

RN SEQUENCE OF 1-25.  
 RP MEDLINE; 8938830.  
 RA DA SILVA N.J., AIRD S.D., SEEBART C., KAISER I.I.:  
 RT "A gyroxin analog from the venom of the bushmaster (Lachesis muta  
 muta).";  
 RL TOXICON 27:763-771(1989).  
 CC -1- FUNCTION: THROMBIN-LIKE SNAKE VENOM SERINE PROTEASE. DISPLAYS A  
 CC SPECIFICITY SIMILAR TO TRYPSIN. RELEASES ONLY FIBRINOPEPTIDE A IN  
 CC THE CONVERSION OF FIBRINOGEN TO FIBRIN.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-GLY BONDS IN  
 CC FIBRINOGEN ALPHA CHAINS (DOES NOT CLEAVE FIBRINOGEN BETA CHAINS).  
 CC -1- INHIBITED COMPETITIVELY BY AMIDINES AND GUANIDINES, AND  
 CC IRREVERSIBLY INHIBITED BY DIISOPROPYLFLUOROPHOSPHATE.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY. BELONGS TO THE SNAKE VENOM SUBFAMILY.  
 DR PIR: A32415; A32415.  
 DR PIR: S35689; S35689.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
 DR PROSITE; PS00135; TRYPSIN\_SER; FALSE\_NEG.  
 DR PFAM; PF00089; trypsin; 1.  
 DR HSP; P00763; IDPO.  
 KW HYDROLASE; SERINE PROTEASE; GLYCOPROTEIN; VENOM.  
 FT ACT\_SITE 43 43 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 88 88 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 177 177 CHARGE RELAY SYSTEM.  
 FT DISULFID 7 138 BY SIMILARITY.  
 FT DISULFID 28 44 BY SIMILARITY.  
 FT DISULFID 78 227 BY SIMILARITY.  
 FT DISULFID 117 183 BY SIMILARITY.  
 FT DISULFID 149 162 BY SIMILARITY.  
 FT DISULFID 173 198 BY SIMILARITY.  
 FT CARBOHYD 45 45  
 FT CARBOHYD 81 81  
 FT CARBOHYD 145 145  
 FT CARBOHYD 224 224  
 SQ SEQUENCE 228 AA; 25629 MW; A0D07DE5 CRC32;  
 Query Match 58.1%; Score 36; DB 1; Length 228;  
 Best Local Similarity 75.0%; Pred. No. 20;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 3 KKKKFFFC 10  
 Db 71 KKKVFFFC 78  
 RESULT 8  
 ID Y410\_BORBU STANDARD; PRT; 39 AA.  
 AC 051371;  
 DT 15-DEC-1998 (REL. 37, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN BB0410.  
 GN BB0410.  
 OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).  
 OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M., CASTENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKY E.K., GINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.:  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 burgdorferi."

RL NATURE 390:580-586(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; AE001146; G2688325; -.  
 DR TIGR; BB0410; -.  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 7 24 POTENTIAL.  
 SQ SEQUENCE 39 AA; 4749 MW; D14379B9 CRC32;  
 Query Match 58.1%; Score 36; DB 1; Length 39;  
 Best Local Similarity 55.6%; Pred. No. 4.3;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Qy 2 KKKKFFFC 10  
 Db 3 KRKFFFLYC 11  
 RESULT 9  
 ID YX10\_BACSU STANDARD; PRT; 428 AA.  
 AC P42306;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 47.3 KD PROTEIN IN WAPA-LICT INTERGENIC REGION.  
 GN YX10 OR S3AR.  
 OS BACILLUS SUBTILIS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 OC BACILLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / BGSCL1;  
 RA YOSHIDA K.-I., SANO H., SEKI S., ODA M., FUJIMURA M., FUJITA Y.;  
 RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).  
 CC -1- SIMILARITY: SOME, TO YEAST YCD8 AND S.POMBE SPAC2G11.13.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; D83026; D1012361; -.  
 DR EMBL; Z99124; E1184635; -.  
 DR SUBTILIST; BG11144; YX10.  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 14 34 POTENTIAL.  
 FT TRANSMEM 55 75 POTENTIAL.  
 FT TRANSMEM 84 104 POTENTIAL.  
 FT TRANSMEM 107 127 POTENTIAL.  
 FT TRANSMEM 149 169 POTENTIAL.  
 FT TRANSMEM 182 202 POTENTIAL.  
 FT TRANSMEM 238 258 POTENTIAL.  
 FT TRANSMEM 272 292 POTENTIAL.  
 FT TRANSMEM 302 322 POTENTIAL.  
 FT TRANSMEM 324 344 POTENTIAL.  
 FT TRANSMEM 361 381 POTENTIAL.  
 FT TRANSMEM 392 412 POTENTIAL.  
 SQ SEQUENCE 428 AA; 47272 MW; 2CA3D0FA CRC32;

```

RESULT 5
MNCP_OXYTR          STANDARD;          PRT;          371 AA.
ID   MNCP_OXYTR
AC   Q27151; 094633;
DT   15-JUL-1998 (REL. 36, CREATED)
DT   15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE   MACRONUCLEAR SOLUTE CARRIER HOMOLOG CR-MSC.
OS   OXYTRICHA TRIFALLAX.
OC   EUKARYOTA; ALVEOLATA; CILIOPHORA; HYPOTRICHS; STICHTOTRICHIDA;
OC   OXYTRICHDIAE; OXYTRICHA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN-JRB310;
RX   MEDLINE; 97109822.
RA   SEG MILLER A., WILLIAMS K.R., HAMMERSMITH R.L., DOAK T.G.,
RA   WITHERSPOON D., MESSICK T., STORJOHANN L.L., HERRICK G.;
RT   "Internal eliminated sequences interrupting the Oxytricha 81 locus:
RT   allelic divergence, conservation, conversions, and possible
RT   transposon origins."
RL   MOL. BIOL. EVOL. 13:1351-1362(1996).
RN   [2]
RP   SEQUENCE OF 4-371 FROM N.A.
RX   MEDLINE; 91367670.
RA   WILLIAMS K.R., HERRICK G.;
RT   "Expression of the gene encoded by a family of macronuclear
RT   chromosomes generated by alternative DNA processing in Oxytricha
RT   fallax."
RL   NUCLEIC ACIDS RES. 19:4717-4724(1991).
CC   -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; L20237; G986927; -
DR   EMBL; M63174; G159761; -
DR   PFAM; PF00153; mito_carr; 2.
KW   TRANSMEMBRANE; TRANSPORT.
FT   TRANSMEM 22 42          POTENTIAL.
FT   TRANSMEM 126 146       POTENTIAL.
FT   TRANSMEM 221 241       POTENTIAL.
FT   TRANSMEM 281 301       POTENTIAL.
FT   CONFLICT 96 96         G -> C (IN REF. 2).
FT   CONFLICT 168 168       V -> I (IN REF. 2).
FT   CONFLICT 262 262       F -> I (IN REF. 2).
FT   CONFLICT 354 354       A -> T (IN REF. 2).
FT   CONFLICT 366 366       D -> E (IN REF. 2).
SQ   SEQUENCE 371 AA; 42547 MW; F92D2D272 CRC32;

Query Match          58.1%; Score 36; DB 1; Length 371;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFF 9
    ||: |||
Db 205 CKNSYEFF 213

RESULT 6
PAMI_YEAST
ID   PAMI_YEAST
AC   P37304;
DT   01-OCT-1994 (REL. 30, CREATED)
DT   01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)

01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
PAMI PROTEIN.
PAMI OR YDR251W OR YD8419.18 OR YD9320A.01.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC   SACCHAROMYCETACEAE; SACCHAROMYCES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN-W303-1A;
RX   MEDLINE; 94148839.
RA   HU G.-Z., RONNE H.;
RT   "Overexpression of yeast PAMI gene permits survival without protein
RT   phosphatase 2A and induces a filamentous phenotype."
RL   J. BIOL. CHEM. 269:3429-3435(1994).
RN   [2]
RP   SEQUENCE OF 660-830 FROM N.A.
RC   STRAIN-S288C / AB972;
RA   MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RA   SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC   -1- FUNCTION: NOT KNOWN. IT IS A SUPPRESSOR OF PROTEIN PHOSPHATASE
CC   2A DEPLETION.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X73454; G456667; -
DR   EMBL; Z70202; E228706; -
DR   EMBL; Z68329; E215297; -
DR   PIR; S45179; S45179.
DR   SGD; L0001334; PAMI.
KW   COILED COIL.
FT   DOMAIN 379 400         COILED COIL.
FT   DOMAIN 481 514         COILED COIL.
FT   DOMAIN 515 522         POLY-GLN.
FT   DOMAIN 810 830         ARG/LYS-RICH (HIGHLY BASIC).
SQ   SEQUENCE 830 AA; 92886 MW; D08F1CD9 CRC32;

Query Match          58.1%; Score 36; DB 1; Length 830;
Best Local Similarity 87.5%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFF 9
    ||||| ||
Db 818 KKKKFSFF 825

RESULT 7
THRL_LACMU
ID   THRL_LACMU          STANDARD;          PRT;          228 AA.
AC   P33589;
DT   01-FEB-1994 (REL. 28, CREATED)
DT   01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT   15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE   GYROXIN ANALOG (EC 3.4.21.74) (THROMBIN-LIKE ENZYME) (VENOMBIN A).
OS   LACHESIS MUTA MUTA (BUSHMASTER).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;
OC   SCLEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CRODALINAE;
OC   LACHESIS.
RN   [1]
RP   SEQUENCE.
RC   TISSUE-VENOM;
RX   MEDLINE; 93359030.
RA   MAGALHAES A., CAMPOS BRASIL DA FONSECA B., RIBEIRO DINIZ C.,
RA   GILROY J., RICHARSON M.;
RT   "The complete amino acid sequence of a thrombin-like enzyme/gyroxin
RT   analogue from venom of the bushmaster snake (Lachesis muta muta)."
RL   FEBS LETT. 329:116-120(1993).

```



RC TISSUE-OVARY;  
RX MEDLINE; 93197888.  
RA FOLTZ K.R., PARTIN J.S., LENNARZ W.J.;  
RT "sea urchin egg receptor for sperm: sequence similarity of binding  
RT domain and hsp70.";  
RL SCIENCE 259:1421-1425(1993).  
CC -!- FUNCTION: CELL SURFACE RECOGNITION PROTEIN THAT BINDS ACROSOME-  
CC REACTED SPERM AND THEREBY MEDIATES BINDING AND SUBSEQUENT FUSION  
CC OF THE SPERM AND EGG.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).  
CC -!- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L04969; G295244; -  
DR PROSITE; PS00297; HSP70.1; FALSE NEG.  
DR PROSITE; PS00329; HSP70.2; FALSE NEG.  
DR PROSITE; PS01036; HSP70.3; 1.  
DR PFAM; PF000012; HSP70; 1.  
DR HSP; P19120; INGC.  
KW SIGNAL; GLYCOPROTEIN; TRANSMEMBRANE; ATP-BINDING.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 1184 EGG SPERM RECEPTOR.  
FT DOMAIN 24 908 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 909 925 POTENTIAL.  
FT DOMAIN 926 1184 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 245 245 POTENTIAL.  
FT CARBOHYD 300 300 POTENTIAL.  
FT CARBOHYD 578 578 POTENTIAL.  
FT CARBOHYD 613 613 POTENTIAL.  
SQ SEQUENCE 1184 AA; 133785 MW; C413034C CRC32;

Query Match 61.3%; Score 38; DB 1; Length 1184;  
Best Local Similarity 50.0%; Pred. No. 41;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
I :|||:  
Db 1080 CFPQKYFFYC 1089

RESULT 3  
YRV9\_CAEEL STANDARD; PRT; 241 AA.  
AC Q09352;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 28.2 KD PROTEIN T10B9.9 IN CHROMOSOME II.  
GN T10B9.9.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; SECERNENTEA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA GARDNER A.;  
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z48717; E1349329; -  
DR WORMPEP; T10B9.9; C601661.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 241 AA; 28172 MW; 857201D8 CRC32;

Query Match 59.7%; Score 37; DB 1; Length 241;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 9  
I :|||:  
Db 28 CKKMFIF 36

RESULT 4  
MNCP\_OXYFA STANDARD; PRT; 371 AA.  
ID MNCP\_OXYFA  
AC P15798; Q94631; O02594; O02506;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MACRONUCLEAR SOLUTE CARRIER HOMOLOG CR-MSC.  
OS OXYTRICHA FALLAX.  
OC EUKARYOTA; ALVEOLATA; CILIOPHORA; HYPOTRICHIS; STICOTRICHIDA;  
OC OXYTRICHIDAE; OXYTRICHA.  
RN [1]  
RP SEQUENCE FROM N.A., AND REVISIONS TO 74-86.  
RX MEDLINE; 97109822.  
RA SEEGMILLER A., WILLIAMS K.R., HAMMERSMITH R.L., DOAK T.G.,  
RA WITHERSPOON D., MESSICK T., STORJOHANN L.L., HERRICK G.;  
RT "Internal eliminated sequences interrupting the Oxytricha 81 locus:  
RT allelic divergence, conservation, conversions, and possible  
RT transposon origins.";  
RT MOL. BIOL. EVOL. 13:1351-1362(1996).  
RN [2]  
RP SEQUENCE OF 74-371 FROM N.A.  
RX MEDLINE; 88112785.  
RA HERRICK G., HUNTER D., WILLIAMS K., KOTTER K.;  
RT "Alternative processing during development of a macronuclear  
RT chromosome family in Oxytricha fallax.";  
RL GENES DEV. 1:1047-1058(1987).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U81495; G2194113; -  
DR PIR; A28260; A28260.  
DR PFAM; PF00153; mito\_carr; 2.  
KW TRANSMEMBRANE; TRANSPORT.  
FT TRANSMEM 22 42 POTENTIAL.  
FT TRANSMEM 126 146 POTENTIAL.  
FT TRANSMEM 221 241 POTENTIAL.  
FT TRANSMEM 281 301 POTENTIAL.  
SQ SEQUENCE 371 AA; 42617 MW; 0E883480 CRC32;

Query Match 58.1%; Score 36; DB 1; Length 371;  
Best Local Similarity 55.6%; Pred. No. 31;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 9  
I :|||:  
Db 205 CKENSYFFC 213

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:21 ; Search time 71.87 Seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-124-280A-19

Perfect score: 62  
Sequence: 1 CKKKKFFFFC 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	38	61.3	1 ADF_MAIZE	P46251 zea mays (m
2	38	61.3	1 ESPR_STRPU	Q06068 strongyloce
3	37	59.7	1 YRV9_CAEEL	Q09352 caenorhabdi
4	36	58.1	1 MNCP_OXYFA	P15798 oxytricha f
5	36	58.1	1 MNCP_OXYTR	Q27151 oxytricha t
6	36	58.1	1 PAM1_YEAST	P37304 saccharomyc
7	36	58.1	1 THRL_LACMO	P33589 lachesis mu
8	36	58.1	1 Y410_BORBU	O51371 borrellia bu
9	36	58.1	1 YX10_BACSU	P42306 bacillus su
10	35	56.5	1 CRL_ECOLI	P24251 escherichia
11	35	56.5	1 YCNX_PORPU	P51364 porphyra pu
12	34	54.8	1 MTH1_HAEIN	P20590 haemophilus
13	34	54.8	1 YH4_YEAST	P38795 saccharomyc
14	34	54.8	1 YVFC_YACCC	P20561 vaccinia vi
15	33	53.2	1 E322_ADEGC	P22229 canine aden
16	33	53.2	1 PAC4_HUMAN	P29122 homo sapien
17	33	53.2	1 PAC4_RAT	Q63415 rattus norv
18	33	53.2	1 SEEB_SALEN	P33387 salmonella
19	33	53.2	1 Y777_METUA	Q58187 methanococc
20	33	53.2	1 YC45_PORPU	P51281 porphyra pu
21	33	53.2	1 YFK5_YEAST	P43608 saccharomyc
22	33	53.2	1 YJ52_YEAST	P45986 saccharomyc
23	33	53.2	1 YW02_MARPO	P38451 marchantia
24	33	53.2	1 YN98_YEAST	P53755 saccharomyc
25	32	51.6	1 ADF_BRANA	P30174 brassica na
26	32	51.6	1 ADF_LILLO	P30175 lilium long
27	32	51.6	1 BCHE_RHOCA	P26168 rhodobacter
28	32	51.6	1 CNJC_TETTH	P16925 tetrahymena
29	32	51.6	1 D7_AEDAE	P18153 aedes aegypt
30	32	51.6	1 ENV_SRV1	P04027 smian retr
31	32	51.6	1 GLNA_HELPY	P94845 helicobacte
32	32	51.6	1 KAZ3_SCHPO	Q09690 schizosacch
33	32	51.6	1 MATK_SOLTU	P32088 solanum tub
34	32	51.6	1 MATK_TOBAC	P12176 nicotiana t
35	32	51.6	1 P2X4_RAT	P51577 rattus norv
36	32	51.6	1 POLG_PVYN	P18247 p genome po
37	32	51.6	1 RM02_PARTE	P15765 paramecium
38	32	51.6	1 RP0D_MARPO	Q05002 human coron
39	32	51.6	1 RPA_CVH22	P06274 marchantia
40	32	51.6	1 SBPI_SARBU	P26228 sarcophaga
41	32	51.6	1 SPCB_DROME	Q00963 drosophila
42	32	51.6	1 SP01_YEAST	P53541 saccharomyc
43	32	51.6	1 Y032_HUMAN	Q15034 homo sapien

ALIGNMENTS

RESULT 1			
ADF_MAIZE			
ID	ADF_MAIZE	STANDARD;	PRT; 139 AA.
AC	P46251.		
DT	01-NOV-1995 (REL. 32, CREATED)		
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)		
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)		
DE	ACTIN DEPOLYMERIZING FACTOR (ADF).		
GN	ABPI.		
OS	ZEAL MAYS (MAIZE).		
OC	EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMERYOPHYTA; TRACHEOPHYTA;		
OC	EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;		
OC	POACEAE; ZEA.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CV. A188; TISSUE-POLLEN;		
RX	MEDLINE; 95232182.		
RA	ROZYCKA M., LOPEZ I., KHAN S., GREENLAND A.J., HUSSEY J.;		
RT	"A zea mays pollen cdna encoding a putative actin-depolymerizing		
RT	factor.";		
RL	PLANT PHYSIOL. 107:1011-1012(1995).		
CC	-!- FUNCTION: ACTIN-DEPOLYMERIZING PROTEIN. SEVERAL ACTIN FILAMENTS		
CC	(F-ACTIN) AND BINDS TO ACTIN MONOMERS (BY SIMILARITY).		
CC	-!- SIMILARITY: BELONGS TO THE ADF FAMILY OF ACTIN-BINDING PROTEINS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL; X80820; G929918; -		
DR	MAIZEDB; 113737; -		
DR	PROSITE; PS00325; ACTIN DEPOLYMERIZING; 1.		
DR	PFAM; PF00241; Cofilin_ADF; 1.		
DR	HSP; Q03048; 1CFY.		
KW	ACTIN-BINDING.		
FT	DOMAIN 92 111 ACTIN-BINDING (POTENTIAL).		
SQ	SEQUENCE 139 AA; 16168 MW; CB6C0B57 CRC32;		
Query Match 61.3% Score 38; DB 1; Length 139;			
Best Local Similarity 66.7%; Pred. No. 6.2;			
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;			
Qy	1 CKKKKFFFF 9		
Db	80 CQKSKIFFF 88		
RESULT 2			
ESPR_STRPU			
ID	ESPR_STRPU	STANDARD;	PRT; 1184 AA.
AC	Q06068.		
DT	01-NOV-1997 (REL. 35, CREATED)		
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)		
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)		
DE	EGG SPERM RECEPTOR PRECURSOR.		
OS	STRONGYLOCENTROTUS PURPURATUS (PURPLE SEA URCHIN).		
OC	EUKARYOTA; METAZOA; ECHINODERMATA; ECHINOZOA; ECHINOIDEA;		
OC	EUECHINOIDEA; ECHINACEA; ECHINOIDA; STRONGYLOCENTROTIDAE;		
OC	STRONGYLOCENTROTUS.		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 174-189 AND 501-515.		

**THIS PAGE BLANK (USPTO)**

C:Genetics:  
A:Genetic code: SCS  
C:Superfamily: ADP,ATP carrier protein repeat homology  
C:Keywords: chromosomal protein

Query Match 58.18; Score 36; DB 2; Length 298;  
Best Local Similarity 55.6%; Pred. No. 39;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CKKKKFFFF 9  
||: :|||  
Db 132 CKENSYFFF 140

Search completed: September 7, 1999, 23:06:33  
Job time: 2481 sec

QY 2 KKKKFFFC 10  
 I: :||:|  
 Db 28 KRFRFFVC 36

RESULT 12  
 B70078  
 conserved hypothetical protein yxio - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998  
 C:Accession: B70078  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni, A.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chodura, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997  
 A:Authors: Fulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidos, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Reger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033  
 A:Accession: B70078  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-428 <RUN>  
 A:Cross-references: GB:299124; GB:AL009126; NID:g2636442; PID:el184635; PID:g2636456  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yxio

Query Match 58.1%; Score 36; DB 2; Length 428;  
 Best Local Similarity 77.8%; Pred. No. 52;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 9  
 I||| I||  
 Db 80 CKKKFFGFF 88

RESULT 13  
 S54547  
 PAM1 protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein YD8419.18; protein YD9320A.01; protein YDR251w  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 08-Jul-1995 #sequence\_revision 23-Aug-1996 #text\_change 06-Feb-1998  
 C:Accession: S54547; A53049; S67308; S67454; S45179  
 R:Oliver, K.; Harris, D.  
 submitted to the EMBL Data Library, May 1995  
 A:Reference number: S54530  
 A:Accession: S54547  
 A:Molecule type: DNA  
 A:Residues: 1-659 <OLI>  
 A:Cross-references: EMBL:249701; NID:g817819; PID:g817837; MIPS:YDR251w  
 R:Hu, G.Z.; Ronne, H.  
 J. Biol. Chem. 269, 3429-3435, 1994  
 A:Title: Overexpression of yeast PAM1 gene permits survival without protein phosphatase  
 A:Reference number: A53049; MUID:94148839  
 A:Accession: A53049  
 A:Molecule type: DNA  
 A:Residues: 1-255, 'T', 257-830 <HUA>  
 A:Cross-references: EMBL:X73454; NID:g456666; PID:g456667  
 R:Murphy, L.; Harris, D.  
 submitted to the EMBL Data Library, December 1995  
 A:Reference number: S61117  
 A:Accession: S67308

A:Molecule type: DNA  
 A:Residues: 660-830 <MUR>  
 A:Cross-references: EMBL:Z68329; NID:g1136205; PID:e215297; PID:g1136206; MIPS:YDR251  
 R:Murphy, L.; Harris, D.  
 submitted to the EMBL Data Library, March 1996  
 A:Reference number: S67454  
 A:Accession: S67454  
 A:Molecule type: DNA  
 A:Residues: 660-830 <MUW>  
 A:Cross-references: EMBL:Z70202; NID:g1226026; PID:e228706; PID:g1226027  
 C:Genetics:  
 A:Gene: SGD:PAM1  
 A:Cross-references: SGD:S0002659; MIPS:YDR251w  
 A:Map position: 4R  
 C:Keywords: coiled coil

Query Match 58.1%; Score 36; DB 2; Length 830;  
 Best Local Similarity 87.5%; Pred. No. 89;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFC 9  
 I||| I||  
 Db 818 KKKKFSFF 825

RESULT 14  
 H71621  
 Ser/Thr protein kinase PFB0150c - malaria parasite (Plasmodium falciparum)  
 C:Species: Plasmodium falciparum  
 C>Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 16-Dec-1998  
 C:Accession: H71621  
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Perte, M.; Salberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
 Science 282, 1126-1132, 1998  
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
 A:Reference number: A71600  
 A:Accession: H71621  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2485 <GAR>  
 A:Cross-references: GB:AE001376; GB:AE001362; NID:g3845108; PID:g3845109; TIGR:PFB015  
 A:Experimental source: clone 3D7  
 C:Genetics:  
 A:Gene: PFB0150c  
 C:Superfamily: protein kinase homology  
 F:2087-2352/Domain: protein kinase homology <KIN>

Query Match 58.1%; Score 36; DB 2; Length 2485;  
 Best Local Similarity 77.8%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 9  
 I||| I||  
 Db 423 CKKKKFSFF 431

RESULT 15  
 A28260  
 Chromosomal protein, macronuclear - Oxytricha fallax (SCC5) (fragment)  
 C:Species: Oxytricha fallax  
 C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 09-Apr-1998  
 C:Accession: A28260  
 R:Herrick, G.; Hunter, D.; Williams, K.; Kotter, K.  
 Genes Dev. 1, 1047-1058, 1987  
 A:Title: Alternative processing during development of a macronuclear chromosome fam1  
 A:Reference number: A28260; MUID:88112785  
 A:Accession: A28260  
 A:Molecule type: DNA  
 A:Residues: 1-298 <HER>  
 A:Cross-references: EMBL:M13033  
 A:Note: the authors translated the codon TAA for residue 96 as Gln and TAG for residu

QY 1 CK----KKKFFFFC 10  
|| | |||||  
DB 79 CKLEVLKSEFFFFC 92

RESULT 7  
G69425  
hypothetical protein AF1408 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
C:Accession: G69425  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343  
A:Accession: G69425  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-144 <KLE>  
A:Cross-references: GB:AE001007; GB:AE000782; NID:g2689330; PID:g2649173; TIGR:AF1408

Query Match 59.7%; Score 37; DB 2; Length 144;  
Best Local Similarity 87.5%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFF 9  
|||||  
DB 41 KKKKVEFF 48

RESULT 8  
S49218  
histidine decarboxylase (EC 4.1.1.22) - Vibrio anguillarum  
C:Species: Vibrio anguillarum  
C>Date: 16-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 05-Dec-1998  
C:Accession: S60898; S49218  
R:Tolmasky, M.E.; Actis, L.A.; Crosa, J.H.  
Mol. Microbiol. 15, 87-95, 1995  
A:Title: A histidine decarboxylase gene encoded by the Vibrio anguillarum plasmid pJM1  
A:Reference number: S60898; MUID:95272396  
A:Accession: S60898  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-386 <TO>  
A:Cross-references: EMBL:Z33880; NID:g535938; PID:g535939  
C:Superfamily: Klebsiella histidine decarboxylase  
C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; pyridoxal phosphate  
F:233/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 58.1%; Score 36; DB 1; Length 386;  
Best Local Similarity 50.0%; Pred. No. 48;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFFC 10  
|| | |||:  
DB 9 CKLNKFWLYC 18

RESULT 9  
S35689  
gyroxin (EC 3.4.21.-) homolog - bushmaster  
C:Species: Lachesis muta muta (bushmaster)  
C>Date: 10-Dec-1993 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
C:Accession: S35689  
R:Nagalhaes, A.; da Fonseca, B.C.B.; Diniz, C.R.; Gilroy, J.; Richardson, M.  
FEBS Lett. 329, 116-120, 1993

A:Title: The complete amino acid sequence of a thrombin-like enzyme/gyroxin analogue  
A:Reference number: S35689; MUID:93359030  
A:Accession: S35689  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-228 <MAG>  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-217/Domain: trypsin homology <TRY>

Query Match 58.1%; Score 36; DB 2; Length 228;  
Best Local Similarity 75.0%; Pred. No. 32;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKKKFFFC 10  
|||||  
DB 71 KKKYFFRC 78

RESULT 10  
A70151  
hypothetical protein BB0410 - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C>Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
C:Accession: A70151  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: A70151  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-39 <KLE>  
A:Cross-references: GB:AE001146; GB:AE000783; NID:g2688312; PID:g2688325; TIGR:BB0410  
A:Experimental source: strain B31

Query Match 58.1%; Score 36; DB 2; Length 39;  
Best Local Similarity 55.6%; Pred. No. 7.9;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KKKKFFFFC 10  
|: |||: |  
DB 3 KRKSKFFLYC 11

RESULT 11  
H71547  
probable disulfide bond chaperone - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C>Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 21-Nov-1998  
C:Accession: H71547  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t  
A:Reference number: A71570; MUID:99000809  
A:Accession: H71547  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-282 <ARN>  
A:Cross-references: GB:AE001291; GB:AE001273; NID:g3328573; PID:g3328581  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: dsbG

Query Match 58.1%; Score 36; DB 2; Length 282;  
Best Local Similarity 55.6%; Pred. No. 38;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of A. thaliana T9A4.

A:Reference number: Z14478  
A:Accession: T01973  
A:Status: translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-162 <ZID>  
A:Cross-references: EMBL:AF096373; NID:g3695400; PID:g3695405  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Map position: 4  
A:Introns: 67/3  
A:Note: T9A4.2

Query Match 62.9%; Score 39; DB 2; Length 162;  
Best Local Similarity 50.0%; Pred. No. 7.9;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CKKKKFFFC 10  
I::I::I::I  
DB 66 CQERKGYFFC 75

RESULT 3  
T02883  
actin depolymerizing factor 2 - maize  
C:Species: Zea mays (maize)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C:Accession: T02883  
R:Lopez, I.; Anthony, R.G.; MacIver, S.K.; Jiang, C.J.; Khan, S.; Weeds, A.G.; Hussey, H.  
Proc. Natl. Acad. Sci. U.S.A. 93, 7415-7420, 1996  
A:Title: Pollen specific expression of maize genes encoding actin depolymerizing factor-  
A:Reference number: Z14757  
A:Accession: T02883  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-139 <LOP>  
A:Cross-references: EMBL:X97725; NID:e993051; PID:e242661  
A:Experimental source: cultivar Al88  
C:Genetics:  
A:Gene: ZmABP2  
C:Keywords: actin binding

Query Match 61.3%; Score 38; DB 2; Length 139;  
Best Local Similarity 66.7%; Pred. No. 10;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CKKKKFFFC 9  
I::I::I::I  
DB 80 CQSKIFFF 88

RESULT 4  
T02882  
actin-depolymerizing factor - maize  
C:Species: Zea mays (maize)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C:Accession: T02882  
R:Rozycka, M.; Lopez, I.; Khan, S.; Greenland, A.J.; Hussey, J.  
Plant Physiol. 107, 1011-1012, 1995  
A:Title: A zein pollen cDNA encoding a putative actin-depolymerizing factor.  
A:Reference number: Z14756  
A:Accession: T02882  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-139 <ROZ>  
A:Cross-references: EMBL:X80820; NID:g929917; PID:g929918  
A:Experimental source: cultivar Al88  
C:Genetics:  
A:Gene: ABP1  
C:Keywords: actin binding

Query Match 61.3%; Score 38; DB 2; Length 139;  
Best Local Similarity 66.7%; Pred. No. 10;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CKKKKFFFC 9  
I::I::I::I  
DB 80 CQSKIFFF 88

RESULT 5  
G71611  
hypothetical protein PFB0560w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C:Accession: G71611  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.  
.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600  
A:Accession: G71611  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2206 <GAR>  
A:Cross-references: GB:AE001403; GB:AE001362; NID:g3845216; PID:g3845217; TIGR:PFB056  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0560w

Query Match 61.3%; Score 38; DB 2; Length 2206;  
Best Local Similarity 87.5%; Pred. No. 90;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 KKKKKFFFC 9  
I::I::I::I  
DB 1285 KKKTKFFFC 1292

RESULT 6  
S69496  
bombyxin B-11 precursor - silkworm  
C:Species: Bombyx mori (silkworm)  
C:Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Mar-1998  
C:Accession: S69496; JQ0837  
R:Kondo, H.; Ino, M.; Suzuki, A.; Ishizaki, H.; Iwami, M.  
J. Mol. Biol. 259, 926-937, 1996  
A:Title: Multiple gene copies for bombyxin, an insulin-related peptide of the silkworm  
cullar forms of bombyxin.  
A:Reference number: S69477; MUID:96266490  
A:Accession: S69496  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-93 <KON>  
A:Cross-references: EMBL:D00788; NID:g217247; PID:d1001142; PID:g217248  
C:Genetics:  
A:Gene: bxb11  
C:Superfamily: insulin  
C:Keywords: pyroglutamic acid  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-46/67-93/Product: bombyxin B-11 #status predicted <MAT>  
F:23-46/Domain: chain B #status predicted <BCH>  
F:49-84/Domain: connecting peptide #status predicted <CPEP>  
F:73-93/Domain: chain A #status predicted <ACH>  
F:23/Modified site: pyroglutamate bonds: #status predicted  
F:30-75,42-92,74-79/Disulfide bonds: #status predicted

Query Match 59.7%; Score 37; DB 2; Length 93;  
Best Local Similarity 57.1%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:32 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-19  
Perfect score: 62  
Sequence: 1 CKKKKFFFC 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	62.9	133	2 A64749	curlin genes trans
2	39	62.9	162	2 T01973	hypothetical prote
3	38	61.3	139	2 T02883	actin depolymerizi
4	38	61.3	139	2 T02882	actin-depolymerizi
5	38	61.3	2206	2 G71611	hypothetical prote
6	37	59.7	93	2 S69496	bombyxin B-11 proc
7	37	59.7	144	2 G69425	hypothetical prote
8	36	58.1	386	1 S42218	histidine decarbox
9	36	58.1	228	2 S35689	gyroxin (EC 3.4.21
10	36	58.1	39	2 A70151	hypothetical prote
11	36	58.1	282	2 H71547	probable disulfide
12	36	58.1	428	2 B70078	conserved hypotet
13	36	58.1	830	2 S54547	PAM1 protein - ya
14	36	58.1	2485	2 H71621	Ser/Thr protein ki
15	36	58.1	298	2 A28260	chromosomal protei
16	35.5	57.3	673	4 A0201	artifact-warnin s
17	35	56.5	98	1 F2NTK	photosystem II pro
18	35	56.5	783	2 T03138	helicase (EC 3.6.1
19	35	56.5	250	2 S62838	hypothetical prote
20	35	56.5	240	2 S73285	hypothetical prote
21	35	56.5	1817	2 D71606	hypothetical prote
22	34	54.8	359	2 J70391	site-specific meth
23	34	54.8	96	2 B42506	F-ORF-C protein -
24	34	54.8	527	2 D71557	probable glucanotr
25	34	54.8	90	2 T03194	hypothetical prote
26	34	54.8	714	2 S46811	hypothetical prote
27	34	54.8	1121	2 F71613	hypothetical prote
28	34	54.8	1351	2 C71607	hypothetical prote
29	34	54.8	368	2 S40627	hypothetical prote
30	33.5	54.0	379	2 F71533	hypothetical prote
31	33	53.2	195	1 ERADD2	early E3 22K glyco
32	33	53.2	969	2 A39490	serine proteinase
33	33	53.2	932	2 I53282	gene PACE4 protei
34	33	53.2	932	2 I53257	PACE4A - mouse (fr
35	33	53.2	296	2 H69796	lactose permease h
36	33	53.2	246	2 B40618	fimbrial periplasm
37	33	53.2	193	2 A64397	hypothetical prote
38	33	53.2	177	2 J70333	conserved hypotet
39	33	53.2	53	2 JS0250	hypothetical 6K pr

ALIGNMENTS

```
RESULT 1
A64749
curlin genes transcription activator - Escherichia coli
N:Alternate names: crl protein
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 29-May-1998
C:Accession: A64749; C22419; S25251; A05115
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: A64749
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-133 <BLAT>
A:Cross-references: GB:AE000132; GB:U00096; NID:q2367098; PID:gl786435; UWGP:b0240
A:Experimental source: strain K-12, substrain MG1655
R:Nuesch, J.; Schumperli, D.
Gene 32, 243-249, 1984
A>Title: Structural and functional organization of the gpt gene region of Escherichia
A:Reference number: A91517; MUID:85155481
A:Accession: C22419
A:Molecule type: DNA
A:Residues: 1-32 <NUC>
A:Cross-references: GB:M13422; NID:q146255; PID:g551809
R:Arngvist, A.; Olsen, A.; Pfeifer, J.; Russell, D.G.; Normark, S.
Mol. Microbiol. 6, 2443-2452, 1992
A>Title: The Crl protein activates cryptic genes for curli formation and fibronectin
A:Reference number: S25251; MUID:93023873
A:Accession: S25251
A:Molecule type: DNA
A:Residues: 1-28,'E',30-133 <ARN>
A:Cross-references: EMBL:X67207; NID:g41153; PID:g41154
A:Experimental source: clinical isolate A012
R:Overbeek, N.; Bergmans, H.; van Mansfeld, F.; Lugtenberg, B.
J. Mol. Biol. 163, 513-532, 1983
A>Title: Complete nucleotide sequence of phoE, the structural gene for the phosphate
A:Reference number: A92893; MUID:83189086
A:Accession: A05115
A:Molecule type: DNA
A:Residues: 'EIA',1-7,'L',9-12,'N',14-133 <OVE>
C:Genetics:
A:Gene: crl

Query Match 62.9%; Score 39; DB 2; Length 133;
Best Local Similarity 60.0%; Pred. No. 6.7;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10
Db 28 CKDNREFFDC 37

RESULT 2
T01973
hypothetical protein T9A4.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse ear cross)
C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999
C:Accession: T01973
R:Zidanic, M.; McQuerry, Y.; Smith, A.
```



**THIS PAGE BLANK (USPTO)**

;; FILING DATE: 12-DEC-1992  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: WO 92/14818  
;; FILING DATE: 12-FEB-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/827,685  
;; FILING DATE: 30-JAN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/657,584  
;; FILING DATE: 19-FEB-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Saliwanchik, David R.  
;; REGISTRATION NUMBER: 31,794  
;; REFERENCE/DOCKET NUMBER: UFI14.C3  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 904-375-8100  
;; TELEFAX: 904-372-5800  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 464 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-07-991-867B-2

Query Match 48.4%; Score 30; DB 1; Length 464;  
Best Local Similarity 71.4%; Pred. No. 7.3e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 KKKEFFF 9  
Db 97 KRKEFLF 103

Search completed: September 7, 1999, 22:38:39  
Job time: 7923 sec

```
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-08-456-112B-42

Query Match 51.6%; Score 32; DB 2; Length 10;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFF 8
Db 3 KKKKFLF 9

RESULT 13
US-07-841-651-4
; Sequence 4, Application US/07841651
; Patent No. 5410031
; GENERAL INFORMATION:
; APPLICANT: Pajor, Ana M
; APPLICANT: Wright, Ernest M
; TITLE OF INVENTION: Cloning and Functional Expression of a
; TITLE OF INVENTION: Mammalian Nat/Nucleoside Cotransporter: A Member of the
; TITLE OF INVENTION: SGLT Family
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,651
; FILING DATE: 19920224
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandel, Saralynn
; REGISTRATION NUMBER: 31,853
; REFERENCE/DOCKET NUMBER: 8772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 662 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Oryctolagus cuniculus
US-07-841-651-4

Query Match 50.0%; Score 31; DB 1; Length 662;
Best Local Similarity 53.3%; Pred. No. 7e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 2 KKKKFF-----FC 10
Db 594 KKKKGFRRAYDLFC 608

RESULT 14
US-08-248-839C-121
; Sequence 121, Application US/08248839C

; Patent No. 5843702
; GENERAL INFORMATION:
; APPLICANT: McConnell, David
; APPLICANT: Devine, Kevin
; APPLICANT: O'Kane, Charles
; TITLE OF INVENTION: A Gene Expression System
; NUMBER OF SEQUENCES: 185
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5843702o No. 5843702disk of No. 5843702th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,839C
; FILING DATE: 25-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta A.
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 3614.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; Molecule Type: Protein
US-08-248-839C-121

Query Match 50.0%; Score 31; DB 2; Length 12;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKKKFF 9
Db 2 RKSFFF 8

RESULT 15
US-07-991-867B-2
; Sequence 2, Application US/07991867B
; Patent No. 5476781
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,867B
```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/02147A  
FILING DATE: 19930309  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,629  
FILING DATE: 09-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: NEELEY PH.D., RICHARD L  
REGISTRATION NUMBER: 30092  
REFERENCE/DOCKET NUMBER: CHIR-009/000US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-7622  
TELEFAX: (415) 857-0663  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 969 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-02147A-2

Query Match 53.2%; Score 33; DB 3; Length 969;  
Best Local Similarity 50.0%; Pred. No. 5e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
Db 952 CERKLFQFC 961

RESULT 11  
US-08-358-160-126  
Sequence 126, Application US/08358160  
Patent No. 5663143  
GENERAL INFORMATION:  
APPLICANT: LEY, Arthur C.  
APPLICANT: LADNER, Robert C.  
APPLICANT: GUTERMAN, Sonia K.  
APPLICANT: ROBERTS, Bruce L.  
APPLICANT: MARKLAND, William  
APPLICANT: KENT, Rachel B.  
TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ  
DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE  
NUMBER OF SEQUENCES: 234  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W. Suite 300  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/358,160  
FILING DATE: 16-DEC-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/133,031  
FILING DATE: 13-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,319

FILING DATE: 26-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/664,989  
FILING DATE: 01-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/487,063  
FILING DATE: 02-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/240,160  
FILING DATE: 02-SEP-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Cooper, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: LEY-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 126:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 57 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-358-160-126

Query Match 51.6%; Score 32; DB 1; Length 57;  
Best Local Similarity 55.6%; Pred. No. 73;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 9  
Db 15 CRKSDFFVF 23

RESULT 12  
US-08-456-112B-42  
Sequence 42, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: PORIO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:

ATTORNEY/AGENT INFORMATION:  
NAME: King, Joshua  
REGISTRATION NUMBER: 35,570  
REFERENCE/DOCKET NUMBER: 920043.403C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836 SEEDANBERRY  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-233-788A-41

Query Match 53.2%; Score 33; DB 1; Length 246;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 FFFFC 10  
| | | | |  
DB 14 FFFFC 18

RESULT 8  
US-08-684-862-5  
Sequence 5, Application US/08684862  
Patent No. 5759541  
GENERAL INFORMATION:  
APPLICANT: Bach, Alfred  
APPLICANT: Hillen, Heinz  
APPLICANT: Bialojan, Siegfried  
TITLE OF INVENTION: No. 5759541el Proteins, the Preparation and Use  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kell & Weinkauff  
STREET: 1101 Connecticut Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage  
COMPUTER: IBM AT-compatible, 80286 processor  
OPERATING SYSTEM: MS-DOS version 5.0  
SOFTWARE: WordPerfect version 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/684,862  
FILING DATE:

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/361,705  
FILING DATE:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,040  
FILING DATE: 30-DEC-1992  
APPLICATION NUMBER: PCT/EP91/01361  
FILING DATE: 19-JUL-1991  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 236 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-684-862-5

Query Match 53.2%; Score 33; DB 2; Length 236;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKKFFFC 10  
| : | : | : |  
DB 71 KEKYFFSC 78  
RESULT 9  
US-08-284-941-2  
Sequence 2, Application US/08284941  
Patent No. 5863756  
GENERAL INFORMATION:  
APPLICANT: BARR, PHILIP J  
APPLICANT: KIEFER, MICHAEL C  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND  
TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM  
STREET: FIVE PALO ALTO SQUARE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94306

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,941  
FILING DATE: 2 August 1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: NEELEY PH.D., RICHARD L.  
REGISTRATION NUMBER: 30092  
REFERENCE/DOCKET NUMBER: CHIR-009/01US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 843-5070  
TELEFAX: (415) 857-0663  
TELEX: 380816 COOLEY PA  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 969 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-941-2

Query Match 53.2%; Score 33; DB 2; Length 969;  
Best Local Similarity 50.0%; Pred. No. 5e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
| : | : | : |  
DB 952 CERKLFIOFC 961

RESULT 10  
PCT-US93-02147A-2  
Sequence 2, Application PC/TUS9302147A  
GENERAL INFORMATION:  
APPLICANT: BARR, PHILIP J  
APPLICANT: KIEFER, MICHAEL C  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND  
TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM  
STREET: FIVE PALO ALTO SQUARE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94306

Query Match 54.8%; Score 34; DB 2; Length 287;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFF 8  
| | | | |  
DB 148 CKKNKFTF 155

RESULT 5  
US-08-471-717-4  
; Sequence 4, Application US/08471717  
; Patent No. 5859337  
; GENERAL INFORMATION:  
; APPLICANT: Gasser, Charles S.  
; APPLICANT: Lippuner, Veronica  
; TITLE OF INVENTION: GENES CONFERRING SALT TOLERANCE AND  
; TITLE OF INVENTION: THEIR USES  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: One Market Plaza, Steuart Street Tower  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105-1492  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,717  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 2307E-606  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 248 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-08-471-717-4

Query Match 54.8%; Score 34; DB 2; Length 248;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
| : : | | |  
DB 60 CQKAAAFIC 69

RESULT 6  
PCT-US95-06556-3  
; Sequence 3, Application PC/TUS9506556  
; GENERAL INFORMATION:  
; APPLICANT: Cully, Doris F.  
; APPLICANT: Arena, Joseph P.  
; APPLICANT: Liu, Ken K.  
; APPLICANT: Vassiliatis, Demetrios  
; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE  
; CHANNELS  
; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wallen, John W.  
; STREET: 126 E. Lincoln Ave., P.O. Box 2000  
; CITY: Rahway  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06556  
; FILING DATE:

; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/249,112  
; FILING DATE: 25-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wallen, John W.  
; REGISTRATION NUMBER: 35,403  
; REFERENCE/DOCKET NUMBER: 19194  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 594-3905  
; TELEFAX: (908) 594-4720  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 510 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-06556-3

Query Match 54.8%; Score 34; DB 3; Length 510;  
Best Local Similarity 55.6%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFF 9  
| | | | |  
DB 479 CKSNKFLFY 487

RESULT 7  
US-08-233-788A-41  
; Sequence 41, Application US/08233788A  
; Patent No. 5635617  
; GENERAL INFORMATION:  
; APPLICANT: Doran, James L.  
; APPLICANT: Kay, William W.  
; APPLICANT: Collinson, Karen S.  
; APPLICANT: Clouthier, Sharon C.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTION  
; OF SALMONELLA  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/233,788A  
; FILING DATE: 26-APR-1994  
; CLASSIFICATION: 435

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;;  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;;  
;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;;  
US-08-456-112B-19

Query Match 100.0%; Score 62; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
|||  
Db 1 CKKKKFFFC 10

RESULT 3  
US-08-249-112-3  
; Sequence 3, Application US/08249112  
; Patent No. 5527703  
; GENERAL INFORMATION:  
; APPLICANT: Cully, Doris F.  
; APPLICANT: Arena, Joseph P.  
; APPLICANT: Liu, Ken K.  
; APPLICANT: Vassiliadis, Demetrios  
; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE  
; TITLE OF INVENTION: CHANNELS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wallen, John W.  
; STREET: 126 E. Lincoln Ave., P.O. Box 2000  
; CITY: Rahway  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/249,112  
; FILING DATE: 25-MAY-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wallen, John W.

;; REGISTRATION NUMBER: 35,403  
;; REFERENCE/DOCKET NUMBER: 19194  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (908) 594-3905  
;; TELEFAX: (908) 594-4720  
;;  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 510 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
US-08-249-112-3  
  
Query Match 54.8%; Score 34; DB 1; Length 510;  
Best Local Similarity 55.6%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFF 9  
|||  
Db 479 CRSNKFY 487

RESULT 4  
US-08-424-641B-10  
; Sequence 10, Application US/08424641B  
; Patent No. 5824523  
; GENERAL INFORMATION:  
; APPLICANT: Sylvain Moineau, Shirley A.  
; APPLICANT: Walker, Ebenezer R. Vedamuthu,  
; APPLICANT: and Peter A. Vandenberg  
; TITLE OF INVENTION: Isolated DNA Encoding  
; TITLE OF INVENTION: Enzyme For Phage  
; TITLE OF INVENTION: Resistance  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ian C. McLeod  
; STREET: 2190 Commons Parkway  
; CITY: Okemos  
; STATE: Michigan  
; COUNTRY: USA  
; ZIP: 48864  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
; MEDIUM TYPE: storage  
; COMPUTER: Acer  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/424,641B  
; FILING DATE: April 19, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/366,480  
; FILING DATE: December 30, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ian C. McLeod  
; REGISTRATION NUMBER: 20,931  
; REFERENCE/DOCKET NUMBER: MT 4.1-151  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (517) 347-4100  
; TELEFAX: (517) 347-4103  
; TELEX: No. 5824523e  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 287 Amino Acids  
; TYPE: Amino Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Peptide  
;;  
US-08-424-641B-10

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:38 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-19

Perfect score: 62

Sequence: 1 CKKKKFFFC 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	10	1	US-08-097-830E-19
2	62	100.0	10	2	US-08-456-112B-19
3	34	54.8	510	1	US-08-249-112-3
4	34	54.8	287	2	US-08-424-641B-10
5	34	54.8	248	2	US-08-471-717-4
6	34	54.8	510	3	PCT-US95-06556-3
7	33	53.2	246	1	US-08-233-788A-41
8	33	53.2	236	2	US-08-684-862-5
9	33	53.2	969	2	US-08-284-941-2
10	33	53.2	969	3	PCT-US93-02147A-2
11	32	51.6	57	1	US-08-358-160-126
12	32	51.6	10	2	US-08-456-112B-42
13	31	50.0	662	1	US-07-841-651-4
14	31	50.0	12	2	US-08-248-839C-121
15	30	48.4	464	1	US-07-991-867B-2
16	30	48.4	365	1	US-08-390-162-2
17	30	48.4	361	1	US-08-390-162-4
18	30	48.4	503	1	US-08-245-294-8
19	30	48.4	503	1	US-08-474-499-8
20	30	48.4	464	1	US-08-107-755A-2
21	30	48.4	372	1	US-08-597-236-8
22	30	48.4	503	1	US-08-307-279A-8
23	30	48.4	385	2	US-08-416-756A-3
24	30	48.4	372	2	US-08-746-682A-8
25	30	48.4	365	2	US-08-685-945B-2
26	30	48.4	361	2	US-08-685-945B-4
27	30	48.4	1706	2	US-08-459-568-2
28	30	48.4	1719	2	US-08-459-568-4
29	30	48.4	86	2	US-08-459-568-82
30	30	48.4	1706	2	US-08-399-411-2
31	30	48.4	1719	2	US-08-399-411-4
32	30	48.4	86	2	US-08-399-411-82
33	30	48.4	417	2	US-08-682-847-6
34	30	48.4	199	2	US-08-405-175A-3
35	30	48.4	200	2	US-08-405-175A-4
36	30	48.4	503	3	PCT-US95-06211-8
37	30	48.4	90	4	5198359-4
38	30	48.4	90	4	5449756-4
39	29	46.8	21	1	US-07-768-269-9

#### ALIGNMENTS

##### RESULT 1

US-08-097-830E-19

; Sequence 19, Application US/08097830E

; Patent No. 5652211

; GENERAL INFORMATION:

; APPLICANT: Porto, Massimo

; TITLE OF INVENTION: Peptides For Neutralizing The

; TITLE OF INVENTION: Toxicity of Lipid A

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson & Costigan, P.C.

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: DOS

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/097,830E

; FILING DATE:

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Costigan, James V.

; REGISTRATION NUMBER: 25,669

; REFERENCE/DOCKET NUMBER: 576-003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 302-8989

; TELEFAX: (212) 302-8998

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10 amino acids

; TYPE: amino acid

; TOPOLOGY: circular

; FEATURE:

; OTHER INFORMATION: sulfide bond between Cys and Cys

US-08-097-830E-19

Query Match

Best Local Similarity 100.0%; Score 62; DB 1; Length 10;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CKKKKFFFC 10

Db 1 CKKKKFFFC 10

##### RESULT 2

US-08-456-112B-19

; Sequence 19, Application US/08456112B

; Patent No. 5834430

; GENERAL INFORMATION:

; APPLICANT: Porto, Massimo

; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

; NUMBER OF SEQUENCES: 45

Sequence 11, Appl  
Sequence 6, Appl  
Sequence 31, Appl  
Sequence 35, Appl  
Sequence 41, Appl  
Sequence 4, Appl



**THIS PAGE BLANK (USPTO)**

ID R94339 standard; peptide; 15 AA.  
 AC R94339;  
 DT 21-AUG-1996 (first entry)  
 DE Anti-thrombogenic peptide.  
 KW Anti-thrombogenic; catheter; blood; tube; bag; dialysing membrane;  
 KW artificial blood vessel; plasma calcium re-addition; inhibitor;  
 KW total blood coagulation; activated partial thromboplastin.  
 OS Synthetic.  
 FH Key  
 FT disulfide\_bond 8..10  
 FT J08059694-A.  
 PN 05-MAR-1996.  
 PD 05-MAR-1996.  
 PF 14-JUN-1995; 172737.  
 PR 15-JUN-1994; JP-158015.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI; 96-184808/19.  
 PT Anti-thrombogenic peptide(s) - useful as anti-thrombogenic agents  
 PT and in stable medical tools with anti-thrombogenic properties  
 PS Example 22; Page 7; lipp; Japanese.  
 CC The peptide is an example of a claimed generic anti-thrombogenic  
 CC (AT) agent, which can be used in stable medical tools with AT  
 CC properties, e.g. catheters, blood tubes, bags and dialysing  
 CC membranes and artificial blood vessels. 15 nm of the peptide when  
 CC added to 100 microl of human plasma, exhibited respective plasma  
 CC calcium re-addn., total blood coagulation and activated partial  
 CC thromboplastin rates of 270, 690 and 68 seconds, while a control  
 CC gp. minus the peptide exhibited corresponding rates of 124, 330  
 CC and 26 seconds.  
 SQ Sequence 15 AA;

Query Match 58.1%; Score 36; DB 1; Length 15;  
 Best Local Similarity 66.7%; Pred. No. 8.2;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KKKKFFFC 10  
 |:::|  
 DB 2 KKKKICFC 10

## RESULT 14

ID P93194 standard; protein; 133 AA.  
 AC P93194;  
 DT 27-MAR-1990 (first entry)  
 DE E.coli fibronectin binding protein designated curlin.  
 KW fnbA gene; fibronectin binding protein; curli pili; curlin.  
 OS Escherichia coli A012.  
 PN EP-342173-A.  
 PD 15-NOV-1989.  
 PF 02-MAY-1989; 850142.  
 PR 06-MAY-1988; SE-001723.  
 PA (NORM/) Normark S.  
 PI Normark S, Olsen A.  
 WPI; 89-334491/46.  
 DR N-PSDB; N92264.  
 PT New fibronectin binding protein expressed as pili on E.coli - and hybrid  
 PT DNA encoding sequences, for preventing E.coli infections by vaccination  
 PT or topical application.  
 PS Disclosure; Fig. 3; 19pp; English.  
 CC The fibronectin binding protein (curlin) is claimed (claim 13, page 12),  
 CC and exists as a polymeric protein on the surface of E.coli in the form of  
 CC curli pili. It has mol.wt. 17 kD, and is expressed by E.coli HB101 (DSM  
 CC 4585) which contains plasmid pFnb20. at 25-30 deg C but not at 37 deg C.  
 CC It is used for immunisation against E.coli infections, eg mastitis or  
 CC urinary or intestinal infections, or to block such infections in open  
 CC skin wounds. When supported on a carrier, it can be used to diagnose  
 CC E.coli infections. For vaccination, dosage is 0.5-5 micrograms/kg, pref.  
 CC 3 doses at 1-3 week intervals, and for topical use an isotonic saline  
 CC soln of concn. 25-250 micrograms FBP/ml is used.  
 SQ Sequence 133 AA;

Query Match 56.5%; Score 35; DB 1; Length 133;  
 Best Local Similarity 50.0%; Pred. No. 71;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
 |:::|  
 DB 28 CEDNREFFDC 37

## RESULT 15

ID R52665 standard; Protein; 133 AA.  
 AC R52665;  
 DT 18-OCT-1994 (first entry)  
 DE Fibronectin binding protein.  
 KW Fibronectin binding protein; FBP; curlin; curli pili; pilus;  
 KW immunization; vaccine; pFnb56.  
 OS Escherichia coli.  
 PN W09410330-A.  
 PD 11-MAY-1994.  
 PR 03-NOV-1993; U10547.  
 PR 03-NOV-1992; US-970846.  
 PA (REAT/) REA T S.  
 PI Normark S, Olsen A;  
 WPI; 94-167484/20.  
 DR P-PSDB; Q62648.  
 PT New fibronectin binding protein curlin from E. coli - used for  
 PT the treatment of infections caused by E. coli and for  
 PT immunisation against E. coli infection  
 PS Disclosure; Fig 3; 33pp; English.  
 CC pFnb10 containing a 3kb large SphI-KpnI fragment expressed  
 CC fibronectin binding when transformed into E. coli HB101. The  
 CC subclone pFnb10 was digested with a series of restriction  
 CC endonucleases to prepare a restriction map. To localize regions  
 CC on pFnb10 expressing fibronectin binding, several subclones  
 CC and deletion derivatives of pFnb10 were constructed and tested for  
 CC their ability to confer binding to E. coli HB101. The sequence  
 CC given in Q62648 is derived from one such construct, and encodes  
 CC a 19 kDa protein.  
 SQ Sequence 133 AA;

Query Match 56.5%; Score 35; DB 1; Length 133;  
 Best Local Similarity 50.0%; Pred. No. 71;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
 |:::|  
 DB 28 CEDNREFFDC 37

Search completed: September 7, 1999, 20:37:21  
 Job time: 18470 sec

CC fatty acids. Phospholipase A2 specifically releases uncommon fatty acids  
 CC from membrane-bound lipids, resulting in their accumulation in the oil  
 CC and preventing the damaging effects of these uncommon fatty acids on cell  
 CC metabolism.  
 SQ Sequence 42 AA;

Query Match 59.7%; Score 37; DB 1; Length 42;  
 Best Local Similarity 50.0%; Pred. No. 14;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
 | : ||: ||  
 Db 23 CEPPKFLFC 32

## RESULT 10

ID R87751 standard; Protein: 140 AA.  
 AC R87751;  
 DT 02-AUG-1996 (first entry)  
 DE MoMLV gag/pol wobble gene product.  
 KW Retrovirus; vector; retrovector; expression cassette;  
 KW gag gene; pol gene; cystic fibrosis; Parkinson disease;  
 KW gene therapy; Moloney murine leukemia virus; MoMLV.  
 OS Synthetic.  
 PN WO9530763-A2.  
 PD 16-NOV-1995.  
 PF 09-MAY-1995; U05789.  
 PR 09-MAY-1994; US-240030.  
 PA (VIAG-) VIAGENE INC.  
 PI Respass JG;  
 DR WPI: 95-404126/51.  
 DR N-PSDB: T08465.  
 PT Retroviral vectors lacking retroviral gag/pol or env coding  
 PT sequences - for producing recombinant viruses for gene therapy of  
 PT e.g. cystic fibrosis, Parkinson disease, etc.  
 PS Example 3B; Page 48; 82pp; English.  
 CC A synthetic gene (T08465) has degenerate codons (i.e. codons that  
 CC 'wobble') for the first 140 amino acids (R87751) of the MoMLV gag  
 CC gene (see also T08452) in addition to synthetic 5' and 3' sequences.  
 CC It is used to produce gag/pol expression cassettes that lack  
 CC non-coding sequences upstream from the gag start, thereby reducing  
 CC recombination potential between the gag/pol expression element and  
 CC the Psi+ sequence of a retroviral vector construct, and inhibiting  
 CC co-packaging of the gag/pol expression element along with the  
 CC retrovirus vector. These vectors are useful for gene therapy.  
 SQ Sequence 140 AA;

Query Match 58.1%; Score 36; DB 1; Length 140;  
 Best Local Similarity 50.0%; Pred. No. 53;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
 | : ||: ||  
 Db 30 CKRRRWTF 39

## RESULT 11

ID R87101 standard; Protein: 140 AA.  
 AC R87101;  
 DT 01-AUG-1996 (first entry)  
 DE Gag gene product.  
 KW Gene delivery vehicle; vector; retrovirus; gene therapy;  
 KW cell targeting; Moloney murine leukaemia virus; MoMLV; gag gene;  
 KW pol gene.  
 OS Synthetic.  
 PN WO9531566-A1.  
 PD 23-NOV-1995.  
 PF 15-MAY-1995; U06084.  
 PR 13-MAY-1994; US-242407.

PA (VIAG-) VIAGENE INC.  
 PI Barber JR, Jolly DJ, Moore MD, Respass JG;  
 DR WPI: 96-010951/01.  
 DR N-PSDB: T07408.  
 PT Targeting of gene delivery vehicles - using a targeting element, a  
 PT high affinity binding pair and a gene delivery vehicle  
 PS Example 3; Page 63-64; 124pp; English.  
 CC A synthetic gene (T07408) has degenerate codons (i.e. codons that  
 CC 'wobble') for the first 140 amino acids (R87101) of the MoMLV gag  
 CC gene (see also T07374) in addition to synthetic 5' and 3' sequences.  
 CC The synthetic gene was used to produce gag/pol expression cassettes  
 CC that lacked non-coding sequences upstream from the gag start, thereby  
 CC reducing recombination potential between the gag/pol expression  
 CC element and the Psi+ sequence of a retroviral vector construct, and  
 CC inhibiting co-packaging of the gag/pol expression element along with  
 CC the retrovirus vector. These vectors are used as gene delivery  
 CC vehicles for targeted gene therapy.  
 SQ Sequence 140 AA;

Query Match 58.1%; Score 36; DB 1; Length 140;  
 Best Local Similarity 50.0%; Pred. No. 53;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
 | : ||: ||  
 Db 30 CKRRRWTF 39

## RESULT 12

ID R94335 standard; peptide: 13 AA.  
 AC R94335;  
 DT 21-AUG-1996 (first entry)  
 DE Anti-thrombogenic peptide.  
 KW Anti-thrombogenic; catheter; blood; tube; bag; dialysing membrane;  
 KW artificial blood vessel; plasma calcium re-addition; inhibitor;  
 KW total blood coagulation; activated partial thromboplastin.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 7..9  
 FN J08059694-A.  
 PD 05-MAR-1996.  
 PF 14-JUN-1995; 172737.  
 PR 15-JUN-1994; JP-158015.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI: 96-184808/19.  
 PT Anti-thrombogenic peptide(s) - useful as anti-thrombogenic agents  
 PT and in stable medical tools with anti-thrombogenic properties  
 PS Example 18; Page 7; 11pp; Japanese.  
 CC The peptide is an example of a claimed generic anti-thrombogenic  
 CC (AT) agent, which can be used in stable medical tools with AT  
 CC properties, e.g. catheters, blood tubes, bags and dialysing  
 CC membranes and artificial blood vessels. 15 nM of the peptide when  
 CC added to 100 microl of human plasma, exhibited respective plasma  
 CC calcium re-addn., total blood coagulation and activated partial  
 CC thromboplastin rates of 285, 720 and 68 seconds, while a control  
 CC gp. minus the peptide exhibited corresponding rates of 124, 330  
 CC and 26 seconds.  
 SQ Sequence 13 AA;

Query Match 58.1%; Score 36; DB 1; Length 13;  
 Best Local Similarity 77.8%; Pred. No. 7.2;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KKKKFFFC 10  
 | : ||: ||  
 Db 1 KKKKFFFC 9

## RESULT 13

R94339

W99556  
 ID W99556 standard; Protein; 352 AA.  
 AC W99556;  
 DT 08-JUN-1999 (first entry)  
 DE Protein encoded by clone CL2 from MSRV-1.  
 KW Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;  
 KW Rheumatoid polyarthritis.  
 OS Multiple sclerosis related virus type 1.  
 PN FR2765588-A1.  
 PD 08-JAN-1999.  
 PF 07-JUL-1997; 008816.  
 PR 07-JUL-1997; FR-008816.  
 PA (INMR ) BIO MERIEUX.  
 DR WPI; 99-098275/09.  
 DR N-PSDB; X29706.  
 PT Nucleic acid sequences of retrovirus called MSRV-1 - associated with  
 PT multiple sclerosis or rheumatoid polyarthritis  
 PS Disclosure; Page 43; 83pp; French.  
 CC This sequence represents the protein encoded by clone CL2 from a  
 CC novel multiple sclerosis related virus type 1 (MSRV1). The sequence  
 CC can be used in diagnostic, prophylactic or therapeutic compositions  
 CC to inhibit expression of a multiple sclerosis related virus and/or  
 CC virus associated with rheumatoid polyarthritis.  
 CC Sequence 352 AA;  
 SQ

Query Match 66.1%; Score 41; DB 1; Length 352;  
 Best Local Similarity 66.7%; Pred. No. 21;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFC 10  
 :||:| |||  
 Db 17 RKKRFFFC 25

RESULT 7  
 W99557  
 ID W99557 standard; Protein; 398 AA.  
 AC W99557;  
 DT 08-JUN-1999 (first entry)  
 DE Protein encoded by pT28C-clone 2 from MSRV-1.  
 KW Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;  
 KW Rheumatoid polyarthritis.  
 OS Multiple sclerosis related virus type 1.  
 PN FR2765588-A1.  
 PD 08-JAN-1999.  
 PF 07-JUL-1997; 008816.  
 PR 07-JUL-1997; FR-008816.  
 PA (INMR ) BIO MERIEUX.  
 DR WPI; 99-098275/09.  
 DR N-PSDB; X29725.  
 PT Nucleic acid sequences of retrovirus called MSRV-1 - associated with  
 PT multiple sclerosis or rheumatoid polyarthritis  
 PS Claim 2; Page 44; 83pp; French.  
 CC This sequence represents the protein encoded by the insert found in  
 CC the plasmid pT28C-clone 2 containing a fragment from a novel multiple  
 CC sclerosis related virus type 1 (MSRV1). The sequence can be used in  
 CC diagnostic, prophylactic or therapeutic compositions to inhibit  
 CC expression of a multiple sclerosis related virus and/or virus  
 CC associated with rheumatoid polyarthritis.  
 CC Sequence 398 AA;  
 SQ

Query Match 66.1%; Score 41; DB 1; Length 398;  
 Best Local Similarity 66.7%; Pred. No. 23;  
 Matches 6; Conservative 2; Mismatches 1; Indels- 0; Gaps 0;

QY 2 KKKKFFFC 10  
 :||:| |||  
 Db 50 RKKRFFFC 58

RESULT 8

W99558  
 ID W99558 standard; Protein; 378 AA.  
 AC W99558;  
 DT 08-JUN-1999 (first entry)  
 DE Protein encoded by pT21C-clone 2 from MSRV-1.  
 KW Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;  
 KW Rheumatoid polyarthritis.  
 OS Multiple sclerosis related virus type 1.  
 PN FR2765588-A1.  
 PD 08-JAN-1999.  
 PF 07-JUL-1997; 008816.  
 PR 07-JUL-1997; FR-008816.  
 PA (INMR ) BIO MERIEUX.  
 DR WPI; 99-098275/09.  
 DR N-PSDB; X29726.  
 PT Nucleic acid sequences of retrovirus called MSRV-1 - associated with  
 PT multiple sclerosis or rheumatoid polyarthritis  
 PS Claim 2; Page 45; 83pp; French.  
 CC This sequence represents the protein encoded by the insert found in  
 CC the plasmid pT21C-clone 2 containing a fragment from a novel multiple  
 CC sclerosis related virus type 1 (MSRV1). The sequence can be used in  
 CC diagnostic, prophylactic or therapeutic compositions to inhibit  
 CC expression of a multiple sclerosis related virus and/or virus  
 CC associated with rheumatoid polyarthritis.  
 CC Sequence 378 AA;  
 SQ

Query Match 66.1%; Score 41; DB 1; Length 378;  
 Best Local Similarity 66.7%; Pred. No. 22;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFC 10  
 :||:| |||  
 Db 30 RKKRFFFC 38

RESULT 9  
 W35381  
 ID W35381 standard; peptide; 42 AA.  
 AC W35381;  
 DT 16-APR-1998 (first entry)  
 DE N-terminal sequence of Ulmus glabra Phospholipase A2.  
 KW Phospholipase A2; distinct acyl specificity; uncommon fatty acid;  
 KW Ulmus glabra.  
 OS Ulmus glabra.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 1 /note= "not specified"  
 FT Misc\_difference 36 /note= "not specified"  
 FT W09737006-A1.  
 PN 09-OCT-1997.  
 PD 27-MAR-1997; SE0554.  
 PR 29-MAR-1996; SE-001237.  
 PA (EKBB/) EK B.  
 PA (SJOE/) SJOE S.  
 PA (STAH/) STAHL U.  
 PA (STVM/) STYME S.  
 PI Ek B, Sjoedahl S, Stahl U, Stymne S;  
 DR WPI; 97-503094/46.  
 PT Phospholipase A2 with specificity for uncommon fatty acids - used to  
 PT produce transgenic oil-producing organisms which accumulate these  
 PT fatty acids for use as chemical feedstocks  
 PS Disclosure; Page 9; 31pp; English.  
 CC The present sequence is represents the N-terminal sequence of the  
 CC Ulmus glabra Phospholipase A2. A novel DNA sequence encodes a low  
 CC molecular weight phospholipase A2, with distinct acyl specificity for  
 CC uncommon fatty acids, that is essentially homologous to Ulmus glabra  
 CC phospholipase A2 or products of the rice cDNA clones D49050, D47724  
 CC and D47653. The present N-terminal sequence was used in a blast search  
 CC program, and the rice cDNA clones identified. The uncommon fatty acids  
 CC are medium chain, very long chain, hydroxy, epoxy, acetylenic acids.  
 CC Phospholipase A2 has specificity for octanoic, decanoic and dodecanoic

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 62; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00084;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
 :||:| |||  
 Db 1 CKKKKFFFC 10

## RESULT 3

W71070  
 ID W71070 standard; Protein: 352 AA.  
 AC W71070;  
 DT 29-DEC-1998 (first entry)  
 DE Multiple sclerosis associated retrovirus protein 5.  
 KW Multiple sclerosis associated retrovirus; MSRV; MS; pol gene;  
 KW gag gene; env gene; rheumatoid arthritis-associated virus.  
 OS Multiple sclerosis associated retrovirus.  
 PN WO9823755-A1.  
 PD 04-JUN-1998.  
 PE 26-NOV-1997; IB1482.  
 PR 26-NOV-1996; US-756429.  
 PA (INMR ) BIO MERIEUX.  
 PI Mandrand B, Paranhos-Baccala G, Perron H;  
 PI WPI; 98-322732/28.  
 DR N-PSDB: V43228.  
 PT New nucleic acid from retroviruses - useful for diagnosis,  
 PT prevention and treatment of, e.g. multiple sclerosis  
 PS Disclosure; Pages 192-194; 286pp; English.  
 CC The present sequence represents a multiple sclerosis (MS) associated  
 CC retrovirus (MSRV) protein used in the method of the invention. The  
 CC invention provides complete or partial genomic sequences of the  
 CC MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by  
 CC these genes. The invention also provides antibodies raised against  
 CC the polypeptides. The genomic sequences, polypeptides and antibodies  
 CC are also claimed useful for diagnosing infection by MS and rheumatoid  
 CC arthritis-associated viruses, and also for prevention and treatment of  
 CC infection with these viruses.  
 SQ Sequence 352 AA;

Query Match 66.1%; Score 41; DB 1; Length 352;  
 Best Local Similarity 66.7%; Pred. No. 21;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFC 10  
 :||:| |||  
 Db 17 RKRRFFFC 25

## RESULT 4

W71071  
 ID W71071 standard; Protein: 398 AA.  
 AC W71071;  
 DT 29-DEC-1998 (first entry)  
 DE Multiple sclerosis associated retrovirus protein 6.  
 KW Multiple sclerosis associated retrovirus; MSRV; MS; pol gene;  
 KW gag gene; env gene; rheumatoid arthritis-associated virus.  
 OS Multiple sclerosis associated retrovirus.  
 PN WO9823755-A1.

PD 04-JUN-1998.  
 PE 26-NOV-1997; IB1482.  
 PR 26-NOV-1996; US-756429.  
 PA (INMR ) BIO MERIEUX.  
 PI Bedin F, Beseme F, Jolivet-Reynaud C, Komurian-Pradel F,  
 PI Mandrand B, Paranhos-Baccala G, Perron H;  
 DR WPI; 98-322732/28.  
 PT New nucleic acid from retroviruses - useful for diagnosis,  
 PT prevention and treatment of, e.g. multiple sclerosis  
 PS Disclosure; Pages 194-196; 286pp; English.  
 CC The present sequence represents a multiple sclerosis (MS) associated  
 CC retrovirus (MSRV) protein used in the method of the invention. The  
 CC invention provides complete or partial genomic sequences of the  
 CC MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by  
 CC these genes. The invention also provides antibodies raised against  
 CC the polypeptides. The genomic sequences, polypeptides and antibodies  
 CC are also claimed useful for diagnosing infection by MS and rheumatoid  
 CC arthritis-associated viruses, and also for prevention and treatment of  
 CC infection with these viruses.  
 SQ Sequence 398 AA;

Query Match 66.1%; Score 41; DB 1; Length 398;  
 Best Local Similarity 66.7%; Pred. No. 23;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFC 10  
 :||:| |||  
 Db 50 RKRRFFFC 58

## RESULT 5

W71072  
 ID W71072 standard; Protein: 378 AA.  
 AC W71072;  
 DT 29-DEC-1998 (first entry)  
 DE Multiple sclerosis associated retrovirus protein 7.  
 KW Multiple sclerosis associated retrovirus; MSRV; MS; pol gene;  
 KW gag gene; env gene; rheumatoid arthritis-associated virus.  
 OS Multiple sclerosis associated retrovirus.  
 PN WO9823755-A1.  
 PD 04-JUN-1998.  
 PE 26-NOV-1997; IB1482.  
 PR 26-NOV-1996; US-756429.  
 PA (INMR ) BIO MERIEUX.  
 PI Bedin F, Beseme F, Jolivet-Reynaud C, Komurian-Pradel F,  
 PI Mandrand B, Paranhos-Baccala G, Perron H;  
 DR WPI; 98-322732/28.  
 PT New nucleic acid from retroviruses - useful for diagnosis,  
 PT prevention and treatment of, e.g. multiple sclerosis  
 PS Disclosure; Pages 196-197; 286pp; English.  
 CC The present sequence represents a multiple sclerosis (MS) associated  
 CC retrovirus (MSRV) protein used in the method of the invention. The  
 CC invention provides complete or partial genomic sequences of the  
 CC MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by  
 CC these genes. The invention also provides antibodies raised against  
 CC the polypeptides. The genomic sequences, polypeptides and antibodies  
 CC are also claimed useful for diagnosing infection by MS and rheumatoid  
 CC arthritis-associated viruses, and also for prevention and treatment of  
 CC infection with these viruses.  
 SQ Sequence 378 AA;

Query Match 66.1%; Score 41; DB 1; Length 378;  
 Best Local Similarity 66.7%; Pred. No. 22;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFC 10  
 :||:| |||  
 Db 30 RKRRFFFC 38

## RESULT 6

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:20 ; Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-19

Perfect score: 62  
Sequence: 1 CKKKKFFFFC 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	10	R71790	Peptide neutralisi
2	62	100.0	10	W21607	Antibiotic potenti
3	41	66.1	352	W71070	Multiple sclerosis
4	41	66.1	398	W71071	Multiple sclerosis
5	41	66.1	378	W71072	Multiple sclerosis
6	41	66.1	352	W99556	Protein encoded by
7	41	66.1	398	W99557	Protein encoded by
8	41	66.1	378	W99558	Protein encoded by
9	37	59.7	42	W35381	N-terminal sequenc
10	36	58.1	140	R87751	MoMIV gag/pol wobb
11	36	58.1	140	R87101	Gag gene product.
12	36	58.1	13	R94335	Anti-thrombogenic
13	36	58.1	13	R94339	Anti-thrombogenic
14	35	56.5	133	P93194	Fibronectin bindin
15	35	56.5	133	R52665	Caenorhabditis ele
16	34	54.8	510	R88360	Arabidopsis STO po
17	34	54.8	248	W01605	Diastereomer pepti
18	34	54.8	27	W35227	Human G-protein co
19	34	54.8	371	W69599	Human 7-transmembr
20	34	54.8	371	W69970	Antipathogenic pep
21	34	54.8	27	W82846	Fibrinogenolytic p
22	33	53.2	236	R20556	Paired basic amino
23	33	53.2	969	R41662	SefB sequence. Eli
24	33	53.2	246	R62753	CORK potassium cha
25	33	53.2	434	R92315	Salmonella enterit
26	33	53.2	246	W23572	N. crassa osp pro
27	33	53.2	1298	W68522	Human secreted pro
28	33	53.2	105	W74887	Human metal respon
29	33	53.2	420	W94670	Gamma glutamyl tra
30	32	51.6	591	R13382	Potyvirus replicas
31	32	51.6	522	R45388	Anti-thrombogenic
32	32	51.6	13	R94331	Rat superior cervi
33	32	51.6	13	R94333	Antibiotic potenti
34	32	51.6	388	W04216	MICFTR based chann
35	32	51.6	10	W21630	MICFTR based chann
36	32	51.6	27	W22827	MICFTR based chann
37	32	51.6	28	W22828	MICFTR based chann
38	32	51.6	29	W22829	MICFTR based chann
39	32	51.6	30	W22830	H. pylori GHPO 114
40	32	51.6	147	W98282	Human p53 regulate
41	32	51.6	388	W78438	Soluble mannose re
42	31	50.0	1456	R24033	Staphylococcal nuc
43	31	50.0	60	R10040	

#### ALIGNMENTS

##### RESULT 1

R71790  
ID R71790 standard; peptide; 10 AA.  
AC R71790;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..10  
PN WO9503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids  
PS Claim 21; Page 21; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which include units of formula: (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new peptides. 10 AA;  
SQ Sequence 10 AA;

Query Match 100.0%; Score 62; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00084;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKKKFFFFC 10  
| | | | | | | | | |  
DB 1 CKKKKFFFFC 10

##### RESULT 2

W21607  
ID W21607 standard; peptide; 10 AA.  
AC W21607;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #19.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide; permeability; outer bacterial membrane.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..10  
PN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M; Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic required  
PS Claim 23; Page 26; 37pp; English.

44 31 50.0 221 1 R10041 Plasmid pOW350 nuc  
45 31 50.0 795 1 W97842 Human P2Y11 recept

**THIS PAGE BLANK (USE)**

QY 1 KCKEKKCK 10  
DB 16 KPRYRKCK 25

## RESULT 14

O29653 PRELIMINARY; PRT; 396 AA.  
AC O29653;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE LPS BIOSYNTHESIS PROTEIN, PUTATIVE.  
GN AF0602.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
OC Archaeoglobus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE; 98049343.  
RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,  
RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRIDES N.C.,  
RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,  
RA OVERBECK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,  
RA COTTON M.D., SPRIGGS T., ARTIACH P., KATNE B.P., SYKES S.M.,  
RA SADOM P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,  
RA VENTER J.C.;  
RT "The complete genome sequence of the hyperthermophilic,  
RT sulphate-reducing archaeon Archaeoglobus fulgidus.";  
RL Nature 390:364-370(1997).  
DR EMBL; AE001063; AAB90634.1; --  
DR TIGR; AF0602; --  
DR PFAM; PF00534; Glycos\_transf\_1; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 396 AA; 45739 MW; DEE352FD CRC32;

Query Match 60.0%; Score 36; DB 1; Length 396;  
Best Local Similarity 66.7%; Pred. NO. 52;  
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 1 KCKEKKCK-CK 10  
DB 83 KAKFKCKCK 94

## RESULT 15

Q50334 PRELIMINARY; PRT; 122 AA.  
AC Q50334;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE D02\_ORF122.  
GN D02\_ORF122.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M129;  
RX MEDLINE; 96177562.  
RA HILBERT H., HIMMELREICH R., PLAGENS H., HERRMANN R.;  
RT "Sequence analysis of 56 kb from the genome of the bacterium  
RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and  
RT a cluster of ribosomal protein genes.";  
RL Nucleic Acids Res. 24:628-639(1996).  
RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN=M129;  
RX MEDLINE; 97105885.  
RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.C.,  
RA HERRMANN R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
RT pneumoniae.";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M129;  
RA HIMMELREICH R., HILBERT H., LI B.C.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U43738; AAC43663.1; --  
DR EMBL; AE000024; AAB95897.1; --  
SQ SEQUENCE 122 AA; 14114 MW; 6C671AB7 CRC32;

Query Match 60.0%; Score 36; DB 2; Length 122;  
Best Local Similarity 60.0%; Pred. NO. 21;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KCKEKKCK 10  
DB 16 KCSFKKCKAE 25

Search completed: September 7, 1999, 20:34:51  
Job time: 19747 sec



KW Envelope protein.  
FT NON\_TER 1  
SQ SEQUENCE 397 AA: 44020 MW: 01EC4EB0 CRC32:

Query Match 61.7%; Score 37; DB 12; Length 397;  
Best Local Similarity 87.5%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KCKFKKFK 8  
||| ||||  
DB 116 KCKDKKFK 123

RESULT 10  
O11548 PRELIMINARY; PRT; 417 AA.  
AC O11548: 1997 (TReMBLrel. 04, Created)  
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BALOTTA C., RIVA C., VIOLIN M., COLOMBO C., GALLI M., MORONI M.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U95451; AAB52798.1; -.  
DR PFAM; PF00516; GP120; 1.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 417  
SQ SEQUENCE 417 AA: 46201 MW: 533FAA5A CRC32:

Query Match 61.7%; Score 37; DB 12; Length 417;  
Best Local Similarity 87.5%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KCKFKKFK 8  
||| ||||  
DB 115 KCKDKKFK 122

RESULT 11  
O11555 PRELIMINARY; PRT; 417 AA.  
AC O11555:  
DT 01-JUL-1997 (TReMBLrel. 04, Created)  
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BALOTTA C., RIVA C., VIOLIN M., COLOMBO C., GALLI M., MORONI M.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U95458; AAB52805.1; -.  
DR PFAM; PF00516; GP120; 1.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 417  
SQ SEQUENCE 417 AA: 46215 MW: 460A3788 CRC32:

Query Match 61.7%; Score 37; DB 12; Length 417;  
Best Local Similarity 87.5%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KCKFKKFK 8  
||| ||||  
DB 115 KCKDKKFK 122

RESULT 12  
O11556 PRELIMINARY; PRT; 413 AA.  
AC O11556:  
DT 01-JUL-1997 (TReMBLrel. 04, Created)  
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BALOTTA C., RIVA C., VIOLIN M., COLOMBO C., GALLI M., MORONI M.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U95459; AAB52806.1; -.  
DR PFAM; PF00516; GP120; 1.  
KW Envelope protein.  
FT NON\_TER 1  
FT NON\_TER 413  
SQ SEQUENCE 413 AA: 45471 MW: A5E4565D CRC32:

Query Match 61.7%; Score 37; DB 12; Length 413;  
Best Local Similarity 87.5%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KCKFKKFK 8  
||| ||||  
DB 119 KCKDKKFK 126

RESULT 13  
Q58145 PRELIMINARY; PRT; 80 AA.  
ID Q58145:  
AC Q58145:  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE RUBREDOXIN.  
GN M30735.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96337999.  
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,  
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L.,  
RA PRESLEY E.A., NGUYEN D., UTTERBACK T.R., KELLEY J.M., PETERSON J.D.,  
RA SADOW P.W., HANNA M.C., COTTON M.D., HURST M.A., ROBERTS K.M.,  
RA KAINE B.P., BORODOVSKY M., KLENK H.P., FRASER C.M., SMITH H.O.,  
RA WOESE C.R., VENTER J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii";  
RL Science 273:1058-1073(1996).  
DR EMBL; U67520; AAB96730.1; -.  
DR PFAM; PF00301; rubredoxin; 1.  
SQ SEQUENCE 80 AA: 9297 MW: 750D8B5B CRC32:

Query Match 60.0%; Score 36; DB 1; Length 80;  
Best Local Similarity 50.0%; Pred. No. 15;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 106 KCKYKFK 112  
|||||

RESULT 6  
Q58160 PRELIMINARY; PRT; 238 AA.  
AC Q58160;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)  
DE HYPOTHETICAL 27.8 KD PROTEIN 0750.  
GN MJ0750.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96337999.  
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,  
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
RA SCOTT J.L., GEORGE N.S.M., WEIDMAN J.F., FUHRMANN J.L.,  
RA PRESLEY E.A., NGUYEN D., UTTERBACK T.R., KELLEY J.M., PETERSON J.D.,  
RA SADOW P.W., HANNA M.C., COTTON M.D., HORST M.A., ROBERTS K.M.,  
RA KATHE B.P., BORODOVSKY M., KLENK H.P., FRASER C.M., SMITH H.O.,  
RA WOESE C.R., VENTER J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii";  
RL Science 273:1058-1073(1996).  
DR EMBL; U67521; AAB96746.1; -;  
DR PROSITE; PS00198; 4FE4S-FERREDOXIN; 1.  
KW Hypothetical protein; Iron-sulfur.  
SQ SEQUENCE 238 AA; 27777 MW; 5261F1D5 CRC32;

Query Match 61.7%; Score 37; DB 1; Length 238;  
Best Local Similarity 66.7%; Pred. NO. 24;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KCKYKFKC 9  
|||||  
Db 190 KCKLCEKFC 198

RESULT 7  
Q94556 PRELIMINARY; PRT; 174 AA.  
AC Q94556;  
DT 01-FEB-1997 (TREMBlrel. 02, Created)  
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE HYPOTHETICAL PROTEIN (FRAGMENT).  
OS Drosophila virilis (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
OC Drosophilidae; Drosophila.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97224612.  
RA ZHOU X., NGUYEN T., KIMBRELL D.A.;  
RT "Identification and characterization of the Cecropin antibacterial  
protein gene locus in Drosophila virilis";  
RL J. Mol. Evol. 44:272-281(1997).  
DR EMBL; U71250; AAB18321.1; -;  
DR FlyBASE; FBgn0025843; Dvir\blastopia.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 174 AA; 19346 MW; C32A8329 CRC32;

Query Match 61.7%; Score 37; DB 5; Length 174;

Best Local Similarity 60.0%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KCKFKFKCK 10  
|||  
Db 24 KCIFKLSCK 33

RESULT 8  
Q95011 PRELIMINARY; PRT; 1162 AA.  
AC Q95011;  
DT 01-FEB-1997 (TREMBlrel. 02, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE F29G6.1 PROTEIN.  
GN F29G6.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA HARRIS B.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans";  
RL Nature 368:32-38(1994).  
CC -!- SIMILARITY: TO OTHER KAZAL TYPE INHIBITORS.  
DR EMBL; Z78543; CAB01753.1; -;  
DR PFAM; PF00050; kazal; 11.  
KW Serine protease inhibitor.  
SQ SEQUENCE 1162 AA; 130179 MW; F0641DA3 CRC32;

Query Match 61.7%; Score 37; DB 5; Length 1162;  
Best Local Similarity 66.7%; Pred. NO. 81;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKFKCK 10  
|:|:|:|  
Db 268 CQFKFFACK 276

RESULT 9  
O11547 PRELIMINARY; PRT; 397 AA.  
AC O11547;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BALOTTA C., RIVA C., VIOLIN M., COLOMBO C., GALLI M., MORONI M.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U95450; AAB52797.1; -;  
DR PFAM; PF00516; GP120; 1.

```

FT NON_TER 349 349
SQ SEQUENCE 349 AA; 38972 MW; 8D58C2A7 CRC32;

Query Match 65.0%; Score 39; DB 12; Length 349;
Best Local Similarity 64.3%; Pred. No. 15;
Matches 9; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

Qy 1 KCKFKFKF----KCK 10
   ||| ||| |||
Db 108 KCKDKKFKSGKCK 121

RESULT 3
P74347
ID P74347 PRELIMINARY; PRT; 243 AA.
AC P74347;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE GALACTOSYL-1-PHOSPHATE TRANSFERASE.
GN RFBP.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=PCC6803;
RA TABATA S.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE; 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res 3:109-136(1996).
DR EMBL; D90914; BAA18441.1; -.
KW Transferase.
SQ SEQUENCE 243 AA; 27788 MW; 46CF5B19 CRC32;

Query Match 63.3%; Score 38; DB 2; Length 243;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KFKRFFKC 9
   |||:||||
Db 88 KFKRFFKC 94

RESULT 4
O65896
ID O65896 PRELIMINARY; PRT; 301 AA.
AC O65896;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE CYTIDINE DEAMINASE (EC 3.5.4.5).
GN CDAL OR CDD OR F3P11.17.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RA FAIVRE-NITSCHKE E.S., GRIENENBERGER J.M., GUALBERTO J.M.;

```

```

RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91156681.
RA ELLEDGE J.E., MULLIGAN J.T., RAMER S.W., SPOTTSHOOD M., DAVIS R.W.;
RT "Lambda YES: a multifunctional cDNA expression vector for the
RT isolation of genes by complementation of yeast and Escherichia coli
RT mutations.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:1731-1745(1991).
RN [3]
RP SEQUENCE FROM N.A.
RA VINCENZETTI S., CAMBI A., NEUHARD J., SCHNORR K., GRELLONI M.,
RA VITA A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CV Columbia;
RA LIN X., KAUL S., SHEA T.P., FUJII C.Y., SHEN M., VANAKEN S.E.,
RA BARNSTEAD M.E., MASON T.M., BOWMAN C.D., RONNING C.M., BENITO M.,
RA CARRERA A.J., CREASY T.H., BUELL C.R., TOWN C.D., NIERMAN W.C.,
RA FRASER C.M., VENTER J.C.;
RT "Arabidopsis thaliana chromosome II BAC F3P11 genomic sequence.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ005687; CAA06671.1; -.
DR EMBL; AJ005261; CAA08460.1; -.
DR EMBL; AC005917; AAD10156.1; -.
KW Hydrolase.
SQ SEQUENCE 301 AA; 32581 MW; 20C9E9AC CRC32;

Query Match 63.3%; Score 38; DB 10; Length 301;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KCKFKFKFC 9
   |||:||||
Db 290 KCEKVFHC 298

RESULT 5
P88241
ID P88241 PRELIMINARY; PRT; 242 AA.
AC P88241;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PATIENT A.
RX MEDLINE; 97184515.
RA McDONALD R.A., MAYERS D.L., CHUNG R.C.Y., WAGNER K.F., KIM S.,
RA BIRX D.L., MICHAEL N.L.;
RT "Evolution of human immunodeficiency virus type 1 env sequence
RT variation in patients with diverse rates of disease progression and
RT T-cell function.";
RL J. Virol. 71:1871-1879(1996).
DR EMBL; U69288; AAC56607.1; -.
DR PFAM; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1 1
FT NON_TER 242 242
SQ SEQUENCE 242 AA; 26863 MW; E3D221BB CRC32;

Query Match 63.3%; Score 38; DB 12; Length 242;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KCKFKFKF 7

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:50 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-18

Perfect score: 60

Sequence: 1 KCKFKKFKCK 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SP TREMBL\_10:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mmc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	65.0	347	12 P87928	P87928 human immun
2	39	65.0	349	12 Q71089	Q71089 human immun
3	38	63.3	243	2 P74347	P74347 synechocyst
4	38	63.3	301	10 O65896	O65896 arabidopsis
5	38	63.3	242	12 P88241	P88241 human immun
6	37	61.7	238	1 Q58160	Q58160 methanococc
7	37	61.7	174	5 Q94556	Q94556 drosophila
8	37	61.7	1162	5 Q95011	Q95011 caenorhabdi
9	37	61.7	397	12 O11547	O11547 human immun
10	37	61.7	417	12 O11548	O11548 human immun
11	37	61.7	417	12 O11555	O11555 human immun
12	37	61.7	413	12 O11556	O11556 human immun
13	36	60.0	80	1 Q58145	Q58145 methanococc
14	36	60.0	396	1 O29653	O29653 archaeoglob
15	36	60.0	122	2 Q50334	Q50334 mycoplasma
16	36	60.0	111	5 Q27069	Q27069 trypanosoma
17	36	60.0	236	11 Q921Y0	Q921Y0 rattus norv
18	35	58.3	326	2 O30713	O30713 flavobacter
19	35	58.3	228	2 O50802	O50802 borrelia bu
20	35	58.3	268	3 Q12286	Q12286 saccharomyc
21	35	58.3	562	4 Q15326	Q15326 homo sapien
22	35	58.3	206	5 Q09659	Q09659 caenorhabdi
23	35	58.3	1296	5 Q22452	Q22452 caenorhabdi
24	35	58.3	279	5 Q18512	Q18512 caenorhabdi
25	35	58.3	455	5 Q19553	Q19553 caenorhabdi
26	35	58.3	1507	5 Q24298	Q24298 drosophila
27	35	58.3	378	12 Q83156	Q83156 mouse cytom
28	35	58.3	359	12 Q71082	Q71082 human immun
29	35	58.3	360	12 Q71087	Q71087 human immun

30	35	58.3	359	12 Q71083	Q71083 human immun
31	35	58.3	856	12 Q73297	Q73297 human immun
32	35	58.3	150	12 Q89783	Q89783 human immun
33	35	58.3	188	13 Q73682	Q73682 brachydanio
34	34	57.5	2588	11 Q88491	Q88491 mus musculu
35	34	56.7	240	1 Q58624	Q58624 methanococc
36	34	56.7	278	2 Q66870	Q66870 aquifex aeo
37	34	56.7	769	3 Q12244	Q12244 saccharomyc
38	34	56.7	557	4 Q00463	Q00463 homo sapien
39	34	56.7	642	5 Q21687	Q21687 caenorhabdi
40	34	56.7	845	5 O18401	O18401 drosophila
41	34	56.7	2910	5 Q26008	Q26008 plasmodium
42	34	56.7	1398	5 Q96244	Q96244 plasmodium
43	34	56.7	1321	5 Q97322	Q97322 plasmodium
44	34	56.7	720	10 Q81072	Q81072 arabidopsis
45	34	56.7	211	12 Q79141	Q79141 human immun

#### ALIGNMENTS

RESULT 1

P87928 PRELIMINARY; PRT; 347 AA.

AC P87928; DT 01-MAY-1997 (Tremblrel. 03, Created)

DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)

DE 01-NOV-1998 (Tremblrel. 08, Last annotation update)

DE ENVELOPE GLYCOPROTEIN, V1-V5 REGION (FRAGMENT).

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retrov. Retroviridae; Lentivirus.

RN [1]

RP SEQUENCE FROM N.A.

RA XIN K.O., SHAPSHAK P.;

RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; U25194; AAB39413.1; -

DR PFAM; PF00516; GP120; 1.

KW Envelope protein.

FT NON\_TER 1

FT NON\_TER 347

SQ SEQUENCE 347 AA; 38285 MW; 32A0F5EE CRC32;

Query Match 65.0%; Score 39; DB 12; Length 347;

Best Local Similarity 64.3%; Pred. No. 15;

Matches 9; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 KCKFKKFK----KCK 10

DB 106 KCKRKFSGKGRCK 119

|||||

|||||

|||||

RESULT 2

Q71089 PRELIMINARY; PRT; 349 AA.

AC Q71089; DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)

DE ENVELOPE GLYCOPROTEIN, V1-V5 REGION (FRAGMENT).

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retrov. Retroviridae; Lentivirus.

RN [1]

RP SEQUENCE FROM N.A.

RA XIN K., SHAPSHAK P., NAGANO I., SRIVASTAVA A.K., WOOD C.,

RA STEWART R.V., MORIKAWA S., USHIJIMA H., OKUDA K.;

RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL; U16091; AAA96679.1; -

DR PFAM; PF00516; GP120; 1.

KW Envelope protein.

FT NON\_TER 1

FT NON\_TER 349

Best Local Similarity 62.5%; Pred. No. 73;  
Matches 5; Conservative 2; Mismatches

1; Indels 0; Gaps 0;

QY 3 KFKKFKCK 10

I: IIII:

Db 219 KYNKFKCR 226

Search completed: September 7, 1999, 23:50:21  
Job time: 1960 sec

PDB; 4PAX; 27-MAY-98.  
 DR PDB; 1A26; 27-MAY-98.  
 DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
 DR PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
 DR PFAM; PF00533; BRCT; 1.  
 DR PFAM; PF00644; PARP; 1.  
 DR PFAM; PF00645; zf-PARP; 2.  
 DR TRANSFERASE; GLYCOSYLTRANSFERASE; NAD; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ADP-RIBOSYLATION; ZINC-FINGER; ZINC; 3D-STRUCTURE.  
 FT DNA\_BIND 1 370  
 FT DOMAIN 371 522  
 FT DOMAIN 523 1011  
 FT ZN\_FING 21 56  
 FT ZN\_FING 125 162  
 FT DOMAIN 207 209  
 FT DOMAIN 220 225  
 FT MOD\_RES 403 404  
 FT MOD\_RES 404 404  
 FT MOD\_RES 410 410  
 FT MOD\_RES 411 411  
 FT MOD\_RES 432 432  
 FT MOD\_RES 434 434  
 FT MOD\_RES 441 441  
 FT MOD\_RES 442 442  
 FT MOD\_RES 453 453  
 FT MOD\_RES 454 454  
 FT MOD\_RES 468 468  
 FT MOD\_RES 481 481  
 FT MOD\_RES 485 485  
 FT MOD\_RES 488 488  
 FT MOD\_RES 509 509  
 FT MOD\_RES 510 510  
 FT MOD\_RES 517 517  
 FT ACT\_SITE 890 890  
 FT CONFLICT 895 895  
 FT SEQUENCE 1011 AA; 113520 MW; 94EE0C16 CRC32;

Query Match 58.3%; Score 35; DB 1; Length 1011;  
 Best Local Similarity 100.0%; Pred. No. 98;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KKEKCK 10  
 |||||  
 DB 345 KKEKCK 350

RESULT 14  
 ID SUC1\_CANAL STANDARD; PRT; 501 AA.  
 AC P3181;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE PROBABLE SUCROSE UTILIZATION PROTEIN SUC1.  
 GN SUC1.  
 OS CANDIDA ALBICANS (YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC CANDIDACEAE; CANDIDA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92104966.  
 RA KELLY R.; KWON-CHUNG K.J.;  
 RT "A zinc finger protein from Candida albicans is involved in sucrose utilization.";  
 RL J. BACTERIOL. 174:222-232(1992).  
 CC -!- FUNCTION: AFFECTS SUCROSE UTILIZATION AND ALPHA-GLUCOSIDASE ACTIVITY. PROBABLE TRANSCRIPTIONAL ACTIVATOR.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR CLUSTER DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).

DR EMBL; S75352; E91548; -  
 DR PIR; A43302; A43302.  
 DR PROSITE; PS00463; ZN2\_CV6\_FUNGAL\_1; 1.  
 DR PROSITE; PS00048; ZN2\_CV6\_FUNGAL\_2; 1.  
 DR PFAM; PF00172; ZN\_Clus; 1.  
 KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;  
 FT ZINC; METAL-BINDING.  
 FT DNA\_BIND 13 39  
 FT SEQUENCE 501 AA; 57110 MW; B445EA95 CRC32;

Query Match 58.3%; Score 35; DB 1; Length 501;  
 Best Local Similarity 62.5%; Pred. No. 55;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFC 9  
 | : | : | : |  
 DB 16 CSFRKVC 23

RESULT 15  
 ID SYTC\_HUMAN STANDARD; PRT; 712 AA.  
 AC P26639;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
 DE THREONYL-TRNA SYNTHETASE, CYTOPLASMIC (EC 6.1.1.3) (THREONINE--TRNA  
 DE LIGASE) (THRRS).  
 GN TARS.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91236775.  
 RA CRUZEN M.E.; ARFIN S.M.;  
 RT "Nucleotide and deduced amino acid sequence of human threonyl-trNA  
 RT synthetase reveals extensive homology to the Escherichia coli and  
 RT yeast enzymes.";  
 RL J. BIOL. CHEM. 266:9919-9923(1991).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-THREONINE + TRNA(THR) = AMP +  
 CC PYROPHOSPHATE + L-THREONYL-TRNA(THR).  
 CC -!- SUBUNIT: HOMODIMER.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).

DR EMBL; M63180; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A38867; YSHUT.  
 DR MIM; 187790; -  
 DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
 DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; 1.  
 DR PFAM; PF00587; trna-synt\_2b; 1.  
 KW AMINOACYL-TRNA SYNTHETASE; PROTEIN BIOSYNTHESIS; LIGASE; ATP-BINDING.  
 FT SEQUENCE 712 AA; 82123 MW; 77B785A5 CRC32;

Query Match 58.3%; Score 35; DB 1; Length 712;

RL FEBS LETT. 276:49-53(1990).  
 RN [8]  
 RP STRUCTURE BY NMR OF ZINC-BINDING REGION.  
 RA MEDLINE: 92212463.  
 RX KRAULIS P.J., RAINE A.R.C., GADHAVI P.L., LAUE E.D.;  
 RT "Structure of the DNA-binding domain of zinc GAL4.";  
 RL NATURE 356:448-450(1992).  
 RN [9]  
 RP STRUCTURE BY NMR OF ZINC-BINDING REGION.  
 RX MEDLINE: 92212464.  
 RA BALEJA J.D., MARWORTH R., HARRISON S.C., WAGNER G.;  
 RT "Solution structure of the DNA-binding domain of Ccd2-GAL4 from S.  
 cerevisiae.";  
 RL NATURE 356:450-453(1992).  
 CC -!- FUNCTION: THIS PROTEIN IS A POSITIVE REGULATOR FOR THE GENE  
 CC EXPRESSION OF THE GALACTOSE-INDUCED GENES SUCH AS GAL1, GAL2,  
 CC GAL7, GAL10, AND MEL1 WHICH ENCODE FOR THE ENZYMES USED TO CONVERT  
 CC GALACTOSE TO GLUCOSE. IT RECOGNIZES A 17 BASE PAIR SEQUENCE IN  
 CC (5'-CGGNNRCYNCCG-3') THE UPSTREAM ACTIVATING SEQUENCE (UAS-G)  
 CC OF THESE GENES.  
 CC -!- SUBUNIT: BINDS DNA AS A HOMODIMER. PROBABLY COMPLEXES WITH GAL11.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- PTM: ASSOCIATION BETWEEN GAL11 AND GAL4 MAY SERVE TO EXPEDITE  
 CC PHOSPHORYLATION OF GAL4.  
 CC -!- THIS PROTEIN CONTAINS ZINC IONS ESSENTIAL FOR DNA BINDING.  
 CC -!- SIMILARITY: STRUCTURALLY AND FUNCTIONALLY RELATED TO LAC9 OF  
 CC KLUYVEROMYCES LACTIS.  
 CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR  
 CC CLUSTER DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: K01486; G171558; .  
 CC DR EMBL: 267751; G1061241; .  
 CC DR EMBL: 273604; E246953; .  
 CC DR PIR: A05022; RCBY64.  
 CC DR PDB: 1D86; 15-APR-93.  
 CC DR PDB: 1AW6; 15-APR-98.  
 CC DR SGD: L0000651; GAL4.  
 CC DR PROSITE: PS00463; ZN2\_CY6\_FUNGAL\_1; 1.  
 CC DR PROSITE: PS00048; ZN2\_CY6\_FUNGAL\_2; 1.  
 CC DR PFAM: PF00172; Zn\_c1us; 1.  
 CC DR TRANSFAC: T00302; .  
 CC KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;  
 CC ZINC; METAL-BINDING; GALACTOSE METABOLISM; PHOSPHORYLATION;  
 CC 3D-STRUCTURE.  
 CC FT DNA\_BIND 11 38 ZN(2)-CYS(6), FUNGAL-TYPE.  
 CC FT MUTAGEN 26 26 P->L: DNA-BINDING LOST.  
 CC FT TURN 12 13  
 CC FT HELIX 14 17  
 CC FT TURN 18 18  
 CC FT TURN 29 30  
 CC FT HELIX 31 34  
 CC FT TURN 35 35  
 CC SQ SEQUENCE 881 AA; 99402 MW; 0D256355 CRC32;

Query Match 58.3%; Score 35; DB 1; Length 881;  
 Best Local Similarity 62.5%; Pred. No. 87;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9  
 Db 14 CRUKKLC 21

RESULT 13

PPOL\_CHICK  
 ID P26446;  
 AC 01-AUG-1992 (REL. 23, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE POLY (ADP-RIBOSE) POLYMERASE (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-  
 DE RIBOSYLTRANSFERASE) (POLY(ADP-RIBOSE) SYNTHETASE).  
 GN ADPRT.  
 OS GALLUS GALLUS (CHICKEN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
 OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OVIDUCT;  
 RX MEDLINE: 91340148.  
 RA ITTEL M.-E., GARNIER J.-M., JELTSCH J.-M., NIEDERGANG C.;  
 RT "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid  
 RT sequence and comparison with mammalian enzyme sequences.";  
 RL GENE 102:157-164(1991).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.  
 RX MEDLINE: 96353841.  
 RA RUF A., MENNISIER DE MURCIA J., DE MURCIA G., SCHULZ G.E.;  
 RT "Structure of the catalytic fragment of poly(AD-ribose) polymerase  
 RT from chicken.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 93:7481-7485(1996).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REV. TO 895.  
 RX MEDLINE: 98191351.  
 RA RUF A., DE MURCIA G., SCHULZ G.E.;  
 RT "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived  
 RT from crystal structures and homology modelling.";  
 RL BIOCHEMISTRY 37:3893-3900(1998).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.  
 RX MEDLINE: 98239716.  
 RA RUF A., ROLLI V., DE MURCIA G., SCHULZ G.E.;  
 RT "The mechanism of the elongation and branching reaction of poly(ADP-  
 RT ribose) polymerase as derived from crystal structures and  
 RT mutagenesis.";  
 RL J. MOL. BIOL. 278:57-65(1998).  
 CC -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
 CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
 CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
 CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
 CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
 CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-RIBOSYL)[N]-ACCEPTOR =  
 CC NICOTINAMIDE + (ADP-D-RIBOSYL)[N+1]-ACCEPTOR.  
 CC -!- COFACTOR: ZINC, CONTAINS TWO MOLE OF ZINC PER MOLE OF PROTEIN.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO AN ACCEPTOR  
 CC CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND FURTHER ADP-  
 CC RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF THE TERMINAL  
 CC ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN AVERAGE CHAIN  
 CC LENGTH OF 20-30 UNITS.  
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X52690; G63743; .  
 CC DR PIR: JH0581; JH0581.  
 CC DR PDB: 2PAW; 27-MAY-98.  
 CC DR PDB: 1PAX; 15-MAY-97.  
 CC DR PDB: 2PAX; 27-MAY-98.  
 CC DR PDB: 3PAX; 27-MAY-98.

PROC. NATL. ACAD. SCI. U.S.A. 89:7432-7436(1992).  
 [2]  
 RL SEQUENCE FROM N.A.  
 RP STRAIN-WISTAR; TISSUE-TESTIS;  
 RC MEDLINE; 91285122.  
 RX MEYERHOF W., MUELLER-BRECHLIN R., RICHTER D.;  
 RA "Molecular cloning of a novel putative G-protein coupled receptor  
 RT expressed during rat spermiogenesis.";  
 RL FEBS LETT. 284:155-160(1991).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE-BRAIN;  
 RC MEDLINE; 96196578.  
 RX SAJJADI F.G., BOYLE D.L., DOMINGO R.C., FIRESTEIN G.S.;  
 RA "CDNA cloning and characterization of A31, an alternatively spliced  
 RT rat A3 adenosine receptor variant.";  
 RL FEBS LETT. 382:125-129(1996).  
 CC -!- FUNCTION: RECEPTOR FOR ADENOSINE. THE ACTIVITY OF THIS RECEPTOR  
 CC IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYL CYCLASE.  
 CC  
 CC -!- POSSIBLE ROLE IN REPRODUCTION.  
 CC  
 CC -!- TISSUE SPECIFICITY: TESTIS, PARTICULARLY IN SPERMATOCYTES AND  
 CC SPERMATIDS BUT NOT IN SPERMATOGENIA. LOW LEVELS IN THE BRAIN.  
 CC  
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED DURING SPERMIOGENESIS.  
 CC  
 CC -!- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY ALTERNATIVE  
 CC SPLICING.  
 CC  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; X59249; G56308; -  
 CC EMBL; M94152; G460332; -  
 CC EMBL; X33219; E213918; -  
 CC PIR; S17177; S17177.  
 CC GCRDB; GCR\_0177; -  
 CC GCRDB; GCR\_0373; -  
 CC GCRDB; GCR\_1362; -  
 CC PROSITE; PS00237; G\_PROTEIN\_RECEPTOR; 1.  
 CC PIR; PF00001; 7tm\_1; 1.  
 CC HSP; P29274; 1MHX.  
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW LIPOPROTEIN; PALMITATE; ALTERNATIVE SPLICING.  
 FT DOMAIN 1 16  
 FT TRANSMEM 17 39  
 FT DOMAIN 40 50  
 FT TRANSMEM 51 74  
 FT DOMAIN 75 86  
 FT TRANSMEM 87 108  
 FT DOMAIN 109 128  
 FT TRANSMEM 129 150  
 FT DOMAIN 151 179  
 FT TRANSMEM 180 200  
 FT DOMAIN 201 233  
 FT TRANSMEM 234 257  
 FT DOMAIN 258 263  
 FT TRANSMEM 264 286  
 FT DOMAIN 287 320  
 FT CARBOHYD 4 4  
 FT CARBOHYD 5 5  
 FT DISULFID 85 168  
 FT LIPID 305 305  
 FT VARSPLIC 119 119  
 FT CONFLICT 18 18  
 FT CONFLICT 68 68  
 FT CONFLICT 74 83  
 FT CONFLICT 132 132  
 F -> L (IN REF. 3).

FT CONFLICT 172 172 S -> F (IN REF. 1).  
 FT CONFLICT 288 288 I -> N (IN REF. 1).  
 FT CONFLICT 291 295 FKTY -> VORNH (IN REF. 1).  
 SQ SEQUENCE 320 AA; 36629 MW; D027EC77 CRC32;  
 Query Match 58.3%; Score 35; DB 1; Length 320;  
 Best Local Similarity 85.7%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 CKFKKFK 8  
 Db 286 CKIKKFK 292  
 RESULT 12  
 GAL4\_YEAST  
 ID GAL4\_YEAST STANDARD; PRT; 881 AA.  
 AC P04386;  
 DT 20-MAR-1987 (REL. 04, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE REGULATORY PROTEIN GAL4.  
 GN GAL4 OR YPL248C.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 CC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 CC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 84141879.  
 RA LAUGHON A., GESTELAND R.F.;  
 RT "Primary structure of the Saccharomyces cerevisiae GAL4 gene.";  
 RL MOL. CELL. BIOL. 4:260-267(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RA POHL T.M.;  
 RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP MUTAGENESIS, AND ZINC REQUIREMENT.  
 RX MEDLINE; 87258276.  
 RA JOHNSTON M.;  
 RT "Genetic evidence that zinc is an essential co-factor in the DNA  
 RT binding domain of GAL4 protein.";  
 RL NATURE 328:353-355(1987).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-65.  
 RX MEDLINE; 92212455.  
 RA MARMORSTEIN R., CAREY M., PTASHNE M., HARRISON S.C.;  
 RT "DNA recognition by GAL4: structure of a protein-DNA complex.";  
 RL NATURE 356:408-414(1992).  
 RN [5]  
 RP STRUCTURE BY NMR OF ZINC-BINDING REGION.  
 RX MEDLINE; 90192750.  
 RA PAN T., COLEMAN J.E.;  
 RT "GAL4 transcription factor is not a 'zinc finger' but forms a  
 RT Zn(II)2Cys6 binuclear cluster.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:2077-2081(1990).  
 RN [6]  
 RP STRUCTURE BY NMR OF ZINC-BINDING REGION.  
 RX MEDLINE; 91214955.  
 RA PAN T., COLEMAN J.E.;  
 RT "Sequential assignments of the 1H NMR resonances of Zn(II)2 and  
 RT 113Cd(II)2 derivatives of the DNA-binding domain of the GAL4  
 RT transcription factor reveal a novel structural motif for specific DNA  
 RT recognition.";  
 RL BIOCHEMISTRY 30:4212-4222(1991).  
 RN [7]  
 RP STRUCTURE BY NMR OF ZINC-BINDING REGION.  
 RX MEDLINE; 91092433.  
 RA GADHAVI P.L., RAINE A.R.C., ALEFOUNDER P.R., LAUE E.D.;  
 RT "Complete assignment of the 1H NMR spectrum and secondary structure  
 RT of the DNA binding domain of GAL4.";



CC -1- FUNCTION: INVOLVED IN DNA REPAIR. MAY HAVE A ROLE IN THE  
CC PROCESSING OF REPLICATION STRUCTURES DURING LATE REPLICATION THAT  
CC IS DIFFERENT FROM ITS ROLE IN THE REPAIR OF RADIATION DAMAGE.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z29640; G452690; -  
CC DR EMBL: Z97208; E325327; -  
CC DR PIR: S41886; S41886;  
CC DR PFAM: PF00176; SNF2\_N; 1.  
CC DR PFAM: PF00271; helicase\_C; 1.  
CC KW DNA REPAIR; NUCLEAR PROTEIN; DNA-BINDING; HELICASE; ATP-BINDING.  
CC FT DOMAIN 35 51 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
CC FT DOMAIN 178 181 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
CC FT DOMAIN 247 251 POLY-PRO.  
CC FT NP\_BIND 294 301 ATP (POTENTIAL).  
CC FT SITE 410 413 DEGH BOX.  
CC FT CONFLICT 261 261 A -> T (IN REF. 1).  
CC SQ SEQUENCE 852 AA; 96653 MW; 170BA3C6 CRC32;

Query Match 60.0%; Score 36; DB 1; Length 852;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FKKFKC 9  
|||||  
DB 34 FKKFKC 39

RESULT 9  
YA06\_CAEEL  
ID YA06\_CAEEL STANDARD; PRT; 1462 AA.  
AC Q20762;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 167.7 KD PROTEIN F54D1.6 IN CHROMOSOME IV.  
GN F54D1.6  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEIA; RHABDITIDA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA LENNARD N.;  
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP REVISIONS.  
RA JONES S.J.M.;  
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- SIMILARITY: STRONG, TO C.ELEGANS R09E10.5.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z77132; E317378; -  
CC DR WORMPEP; F54D1.6; CE11060.  
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
FT TRANSMEM 1346 1366 POTENTIAL.

SQ SEQUENCE 1462 AA; 167666 MW; 5ED4E58A CRC32;  
Query Match 60.0%; Score 36; DB 1; Length 1462;  
Best Local Similarity 55.6%; Pred. No. 92;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 CKFKKFKCK 10  
|||||  
DB 828 CKFADFRCQ 836  
RESULT 10  
YF1M\_CAEEL  
ID YF1M\_CAEEL STANDARD; PRT; 1459 AA.  
AC Q21874;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 166.0 KD PROTEIN R09E10.5 IN CHROMOSOME IV.  
GN R09E10.5  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEIA; RHABDITIDA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA MATTHEWS L.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- SIMILARITY: STRONG, TO C.ELEGANS F54D1.6.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z70287; E1348585; -  
CC DR WORMPEP; R09E10.5; CE06287.  
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
SQ SEQUENCE 1459 AA; 165994 MW; 46E17445 CRC32;  
Query Match 60.0%; Score 36; DB 1; Length 1459;  
Best Local Similarity 55.6%; Pred. No. 92;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 CKFKKFKCK 10  
|||||  
DB 837 CKFADFRCQ 845  
RESULT 11  
AA3R\_RAT  
ID AA3R\_RAT STANDARD; PRT; 320 AA.  
AC P28647; O63792;  
DT 01-DEC-1992 (REL. 24, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE ADENOSINE A3 RECEPTOR (TGPCR1).  
GN ADORA3.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC TISSUE-BRAIN;  
RX MEDLINE; 92366475.  
RA ZHOU Q.Y., LI C.Y., OLAH M.E., JOHNSON R.A., STILES G.L., CIVELLI O.;  
RT "Molecular cloning and characterization of an adenosine receptor: the  
RT A3 adenosine receptor."

```

FT DOMAIN 2900 2938 LDL-RECEPTOR CLASS A 20. BY SIMILARITY.
FT DOMAIN 2939 EGF-LIKE 11. BY SIMILARITY.
FT DOMAIN 2979 EGF-LIKE 12. BY SIMILARITY.
FT DOMAIN 3287 EGF-LIKE 13. BY SIMILARITY.
FT DOMAIN 3329 LDL-RECEPTOR CLASS A 21. BY SIMILARITY.
FT DOMAIN 3369 LDL-RECEPTOR CLASS A 22. BY SIMILARITY.
FT DOMAIN 3407 LDL-RECEPTOR CLASS A 23. BY SIMILARITY.
FT DOMAIN 3447 LDL-RECEPTOR CLASS A 24. BY SIMILARITY.
FT DOMAIN 3488 LDL-RECEPTOR CLASS A 25. BY SIMILARITY.
FT DOMAIN 3489 LDL-RECEPTOR CLASS A 26. BY SIMILARITY.
FT DOMAIN 3531 LDL-RECEPTOR CLASS A 27. BY SIMILARITY.
FT DOMAIN 3569 LDL-RECEPTOR CLASS A 28. BY SIMILARITY.
FT DOMAIN 3570 LDL-RECEPTOR CLASS A 29. BY SIMILARITY.
FT DOMAIN 3608 LDL-RECEPTOR CLASS A 30. BY SIMILARITY.
FT DOMAIN 3646 LDL-RECEPTOR CLASS A 31. BY SIMILARITY.
FT DOMAIN 3689 EGF-LIKE 14. BY SIMILARITY.
FT DOMAIN 3736 EGF-LIKE 15. BY SIMILARITY.
FT DOMAIN 3779 EGF-LIKE 16. BY SIMILARITY.
FT DOMAIN 3821 EGF-LIKE 17. BY SIMILARITY.
FT DOMAIN 3859 EGF-LIKE 18. BY SIMILARITY.
FT DOMAIN 4182 EGF-LIKE 19. BY SIMILARITY.
FT DOMAIN 4231 EGF-LIKE 20. BY SIMILARITY.
FT DOMAIN 4267 EGF-LIKE 21. BY SIMILARITY.
FT DOMAIN 4303 EGF-LIKE 22. BY SIMILARITY.
FT DOMAIN 4339 RECOGNITION SITE FOR PROTEOLYTICAL
FT DOMAIN 4372 PROCESSING (POTENTIAL).
FT SITE 3939 CRITICAL FOR ENDOCYTOSIS (BY SIMILARITY).
FT SITE 4472 CRITICAL FOR ENDOCYTOSIS (BY SIMILARITY).
FT SITE 4506 BY SIMILARITY.
FT SITE 4506 BY SIMILARITY.
FT SITE 29 42 BY SIMILARITY.
FT SITE 36 55 BY SIMILARITY.
FT SITE 49 66 BY SIMILARITY.
FT SITE 74 87 BY SIMILARITY.
FT SITE 81 100 BY SIMILARITY.
FT SITE 110 126 BY SIMILARITY.
FT SITE 122 135 BY SIMILARITY.
FT SITE 137 150 BY SIMILARITY.
FT SITE 156 166 BY SIMILARITY.
FT SITE 162 175 BY SIMILARITY.
FT SITE 177 190 BY SIMILARITY.
FT SITE 480 495 BY SIMILARITY.
FT SITE 491 506 BY SIMILARITY.
FT SITE 508 521 BY SIMILARITY.
FT SITE 805 816 BY SIMILARITY.
FT SITE 812 825 BY SIMILARITY.
FT SITE 827 840 BY SIMILARITY.
FT SITE 852 864 BY SIMILARITY.
FT SITE 859 877 BY SIMILARITY.
FT SITE 871 888 BY SIMILARITY.
FT SITE 893 905 BY SIMILARITY.
FT SITE 900 918 BY SIMILARITY.
FT SITE 912 929 BY SIMILARITY.
FT SITE 934 946 BY SIMILARITY.
FT SITE 941 959 BY SIMILARITY.
FT SITE 953 969 BY SIMILARITY.
FT SITE 974 987 BY SIMILARITY.
FT SITE 982 1000 BY SIMILARITY.
FT SITE 994 1009 BY SIMILARITY.
FT SITE 1013 1025 BY SIMILARITY.
FT SITE 1020 1038 BY SIMILARITY.
FT SITE 1032 1049 BY SIMILARITY.
FT SITE 1060 1073 BY SIMILARITY.
FT SITE 1067 1086 BY SIMILARITY.
FT SITE 1080 1095 BY SIMILARITY.
FT SITE 1102 1116 BY SIMILARITY.
FT SITE 1110 1129 BY SIMILARITY.
FT SITE 1123 1138 BY SIMILARITY.
FT SITE 1143 1157 BY SIMILARITY.
FT SITE 1150 1170 BY SIMILARITY.
FT SITE 1164 1180 BY SIMILARITY.
FT SITE 1183 1194 BY SIMILARITY.
FT SITE 1190 1204 BY SIMILARITY.
FT SITE 1206 1219 BY SIMILARITY.
FT DISULFID 1225 1235 BY SIMILARITY.
FT DISULFID 1231 1244 BY SIMILARITY.
FT DISULFID 1246 1259 BY SIMILARITY.
FT DISULFID 1538 1551 BY SIMILARITY.
FT DISULFID 1547 1561 BY SIMILARITY.
FT DISULFID 1563 1576 BY SIMILARITY.
FT DISULFID 1846 1857 BY SIMILARITY.
FT DISULFID 1853 1867 BY SIMILARITY.
FT DISULFID 1869 1882 BY SIMILARITY.
FT DISULFID 2155 2166 BY SIMILARITY.
FT DISULFID 2162 2176 BY SIMILARITY.
FT DISULFID 2178 2190 BY SIMILARITY.
FT DISULFID 2476 2487 BY SIMILARITY.
FT DISULFID 2483 2497 BY SIMILARITY.
FT DISULFID 2499 2511 BY SIMILARITY.
FT DISULFID 2518 2531 BY SIMILARITY.
FT DISULFID 2526 2544 BY SIMILARITY.
FT DISULFID 2538 2555 BY SIMILARITY.
FT DISULFID 2560 2572 BY SIMILARITY.
FT DISULFID 2567 2585 BY SIMILARITY.
FT DISULFID 2579 2594 BY SIMILARITY.
FT DISULFID 2599 2611 BY SIMILARITY.
FT DISULFID 2606 2624 BY SIMILARITY.
FT DISULFID 2618 2633 BY SIMILARITY.
FT DISULFID 2638 2660 BY SIMILARITY.
FT DISULFID 2654 2673 BY SIMILARITY.
FT DISULFID 2667 2682 BY SIMILARITY.
FT DISULFID 2690 2702 BY SIMILARITY.
FT DISULFID 2697 2715 BY SIMILARITY.
FT DISULFID 2709 2724 BY SIMILARITY.
FT DISULFID 2732 2744 BY SIMILARITY.
FT DISULFID 2739 2757 BY SIMILARITY.
FT DISULFID 2751 2767 BY SIMILARITY.
FT DISULFID 2772 2785 BY SIMILARITY.
FT DISULFID 2779 2798 BY SIMILARITY.
FT DISULFID 2792 2810 BY SIMILARITY.

Query Match 60.0%; Score 36; DB 1: Length 4543;
Best Local Similarity 40.0%; Pred No. 2.3e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 KCKFKFKCK 10
Db 3609 RCEFDQYCK 3618

RESULT 8
RA54_SCHPO
ID RA54_SCHPO STANDARD; PRT: 852 AA.
AC P41410; O13723;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE DNA REPAIR PROTEIN RHP54.
GN RHP54 OR RAD54 OR SPAC15A10.03C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETES;
OC SCHIZOSACCHAROMYCETES; SCHIZOSACCHAROMYCETACEAE;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96431710.
RA MURIS D.F.R., VREEKEN K., CARR A.M., MURRAY J.M., SMIT C.,
FT LOHMAN P.H.M., PASTINK A.;
RT "Isolation of the Schizosaccharomyces pombe RAD54 homologue, rhp54+,
a gene involved in the repair of radiation damage and replication
fidelity.";
RL J. CELL SCI. 109:73-81(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA MURPHY L., HARRIS D., WOOD V., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.

```

CC SERINE- AND THREONINE-SPECIFIC ENZYME. THIS ISOZYME IS A NEGATIVE  
 CC REGULATOR OF THE VISUAL TRANSDUCTION CASCADE AND HAS BEEN SHOWN TO  
 CC BE REQUIRED FOR PHOTORECEPTOR CELL INACTIVATION AND LIGHT  
 CC ADAPTATION.  
 CC -|- SIMILARITY: EXCLUSIVELY EXPRESSED IN PHOTORECEPTOR CELLS.  
 CC -|- SIMILARITY: CONTAINS TWO COPIES OF THE ZINC-DEPENDENT PHORBOL-  
 CC ESTER AND DAG BINDING DOMAIN.  
 CC -|- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J04845; GI58127; -.  
 CC PIR: A32392; A32392.  
 CC FLYBASE: FBgn0004784; InaC.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC PROSITE: PS00479; DAG\_PE\_BINDING\_DOMAIN; 2.  
 CC PROSITE: PS00499; C2\_DOMAIN\_1; 1.  
 CC PROSITE: PS00004; C2\_DOMAIN\_2; 1.  
 CC PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 CC PFAM: PF00069; pkinase; 1.  
 CC PFAM: PF00130; DAG\_PE\_bind; 2.  
 CC PFAM: PF00168; C2; 1.  
 CC PFAM: PF00433; pkinase\_C; 1.  
 CC HSP: P04410; I425.  
 CC  
 CC CALCIUM-BINDING; DUPLICATION; ATP-BINDING; TRANSFERASE;  
 CC SERINE/THREONINE-PROTEIN KINASE; MULTIGENE FAMILY; VISION; ZINC;  
 CC PHORBOL-ESTER BINDING.  
 CC DOMAIN 72 121 PHORBOL-ESTER AND DAG BINDING.  
 CC DOMAIN 137 186 PHORBOL-ESTER AND DAG BINDING.  
 CC DOMAIN 206 269 C2 DOMAIN.  
 CC DOMAIN 371 629 PROTEIN KINASE.  
 CC NP\_BIND 377 385 ATP (BY SIMILARITY).  
 CC BINDING 400 400 ATP (BY SIMILARITY).  
 CC ACT\_SITE 495 495 BY SIMILARITY.  
 CC SEQUENCE 700 AA; 79843 MW; 51DD55CE CRC32;  
 CC  
 CC Query Match 60.0%; Score 36; DB 1; Length 700;  
 CC Best Local Similarity 75.0%; Pred. No. 50;  
 CC Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 CC Qy 2 CKFKKFC 9  
 CC ||| |||  
 CC Db 114 CKFVVFRC 121  
 CC  
 CC RESULT 7  
 CC ID LRP1\_CHICK STANDARD; PRT: 4543 AA.  
 CC AC P98157;  
 CC DT 01-OCT-1996 (REL. 34, CREATED)  
 CC DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 CC DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 CC DE LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 1 PRECURSOR (LRP)  
 CC DE (ALPHA-2-MACROGLOBULIN RECEPTOR) (AZMR).  
 CC OS GALLUS GALLUS (CHICKEN).  
 CC EKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
 CC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
 CC [1]  
 CC RN SEQUENCE FROM N.A.  
 CC RC STRAIN-WHITE LEHIGH; TISSUE=LIVER, AND OVARY;  
 CC RA MEDLINE: 94103212.  
 CC RX NIMPE J., STIFANI S., BILLOUS P.T., SCHNEIDER W.J.;  
 CC "The somatic cell-specific low density lipoprotein receptor-related  
 CC protein of the chicken. Close kinship to mammalian low density  
 CC lipoprotein receptor gene family members.";

RL J. BIOL. CHEM. 269:212-219(1994).  
 CC -|- FUNCTION: INVOLVED IN THE PLASMA CLEARANCE OF CHYLOMICRON REMNANTS  
 CC AND ACTIVATED ALPHA 2-MACROGLOBULIN, AS WELL AS THE LOCAL  
 CC METABOLISM OF COMPLEXES BETWEEN PLASMINOGEN ACTIVATORS AND THEIR  
 CC ENDOGENOUS INHIBITORS. BINDS VITELLOGENIN, CALCIUM AND ALPHA 2-  
 CC MACROGLOBULIN.  
 CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -|- TISSUE SPECIFICITY: SOMATIC.  
 CC -|- PTM: CLEAVED INTO A 85 KD MEMBRANE-SPANNING SUBUNIT (LRP-85) AND  
 CC A 515 KD LARGE EXTRACELLULAR DOMAIN (LRP-515) THAT REMAINS NON-  
 CC COVALENTLY ASSOCIATED.  
 CC -|- ALTERNATIVE PRODUCTS: IN CLONE JN18, AN ASP IS REPLACED BY  
 CC SER-GLU-ARG-GLN-ASP DUE TO ALTERNATIVE SPLICING OF EXON3.  
 CC -|- SIMILARITY: CONTAINS 22 EGF-LIKE DOMAINS.  
 CC -|- SIMILARITY: CONTAINS 31 LDL-RECEPTOR CLASS A DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X74904; G438007; -.  
 CC PROSITE: PS00010; ASX\_HYDROXYL; 3.  
 CC PROSITE: PS00022; EGF\_1; 5.  
 CC PROSITE: PS01186; EGF\_2; 7.  
 CC PROSITE: PS01187; EGF\_CA; 2.  
 CC PROSITE: PS01209; LDLRA\_1; 27.  
 CC PROSITE: PS00068; LDLRA\_2; 31.  
 CC PFAM: PF00008; EGF; 14.  
 CC PFAM: PF00057; ldl\_recept\_a; 31.  
 CC HSP: P01130; LAJ1.  
 CC RECEPTOR; TRANSMEMBRANE; REPEAT; ENDOCYTOSIS; GLYCOPROTEIN;  
 CC SIGNAL; CALCIUM-BINDING; EGF-LIKE DOMAIN; COATED PITS;  
 CC ALTERNATIVE SPLICING.  
 CC SIGNAL 1 21 POTENTIAL.  
 CC CHAIN 22 4543 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED  
 CC PROTEIN 1.  
 CC DOMAIN 22 4419 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 4420 4443 POTENTIAL.  
 CC DOMAIN 4444 4543 CYTOPLASMIC (POTENTIAL).  
 CC DOMAIN 27 68 LDL-RECEPTOR CLASS A 1.  
 CC DOMAIN 72 112 LDL-RECEPTOR CLASS A 2.  
 CC DOMAIN 113 151 EGF-LIKE 1.  
 CC DOMAIN 152 191 EGF-LIKE 2.  
 CC DOMAIN 476 522 EGF-LIKE 3.  
 CC DOMAIN 801 841 EGF-LIKE 4.  
 CC DOMAIN 850 890 LDL-RECEPTOR CLASS A 3.  
 CC DOMAIN 891 931 LDL-RECEPTOR CLASS A 4.  
 CC DOMAIN 932 971 LDL-RECEPTOR CLASS A 5.  
 CC DOMAIN 972 1011 LDL-RECEPTOR CLASS A 6.  
 CC DOMAIN 1011 1051 LDL-RECEPTOR CLASS A 7.  
 CC DOMAIN 1058 1097 LDL-RECEPTOR CLASS A 8.  
 CC DOMAIN 1100 1140 LDL-RECEPTOR CLASS A 9.  
 CC DOMAIN 1141 1180 LDL-RECEPTOR CLASS A 10.  
 CC DOMAIN 1181 1220 EGF-LIKE 5.  
 CC DOMAIN 1221 1260 EGF-LIKE 6.  
 CC DOMAIN 1534 1577 EGF-LIKE 7.  
 CC DOMAIN 1842 1883 EGF-LIKE 8.  
 CC DOMAIN 2151 2191 EGF-LIKE 9.  
 CC DOMAIN 2472 2512 EGF-LIKE 10.  
 CC DOMAIN 2516 2557 LDL-RECEPTOR CLASS A 11.  
 CC DOMAIN 2558 2596 LDL-RECEPTOR CLASS A 12.  
 CC DOMAIN 2597 2635 LDL-RECEPTOR CLASS A 13.  
 CC DOMAIN 2636 2684 LDL-RECEPTOR CLASS A 14.  
 CC DOMAIN 2688 2730 LDL-RECEPTOR CLASS A 15.  
 CC DOMAIN 2730 2769 LDL-RECEPTOR CLASS A 16.  
 CC DOMAIN 2770 2812 LDL-RECEPTOR CLASS A 17.  
 CC DOMAIN 2814 2853 LDL-RECEPTOR CLASS A 18.  
 CC DOMAIN 2854 2897 LDL-RECEPTOR CLASS A 19.

RT and codon usage.":  
 J. MOL. BIOL. 180:239-250(1984).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RA VANDENBOL M., PORTELETT D., HILGER F.;  
 [3]  
 RN SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RP SEQUENCE OF 1-513 FROM N.A.  
 RC STRAIN-S288C;  
 RA SAVILLE S.P., ATKINSON S., JAMIESON L., POCKLINGTON M.J., ORR E.,  
 LEVIN D.E., STEVENSON W.D.;  
 [4]  
 RN SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF 29-123.  
 RX MEDLINE; 95047390.  
 RA MARMORSTEIN R., HARRISON S.C.;  
 RT "Crystal structure of a PP1-DNA complex: DNA recognition by proteins  
 containing a Zn2Cys6 binuclear cluster.";  
 RL GENES DEV. 8:2504-2512(1994).  
 CC -!- FUNCTION: POSITIVE REGULATOR OF URA1 AND URA3 EXPRESSION.  
 CC -!- SUBUNIT: BINDS DNA AS A HOMODIMER.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR  
 CLUSTER DOMAIN.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 -----  
 CC EMBL; X01739; G4215; -  
 DR EMBL; Z73186; E24570; -  
 DR EMBL; X90564; G975228; -  
 DR PIR; S05877; RCBYPI  
 DR PDB; 1PTI; 27-FEB-95.  
 DR SGD; L0001475; PP1.  
 DR PROSITE; PS00463; ZN2\_CYS6\_FUNGAL\_1; 1.  
 DR PROSITE; PS00463; ZN2\_CYS6\_FUNGAL\_2; 1.  
 DR PFAM; PF00172; Zn.clus; 1.  
 DR HSP; P04386; 1AW6.  
 DR TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC; METAL-BINDING; PYRIMIDINE BIOSYNTHESIS; 3D-STRUCTURE.  
 FT DNA\_BIND 34 61 ZN(2)-CYS(6), FUNGAL-TYPE.  
 SQ SEQUENCE 904 AA; 102723 MW; 8C7CF100 CRC32;  
 -----  
 Query Match 61.7%; Score 37; DB 1; Length 904;  
 Best Local Similarity 55.6%; Pred. No. 43;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KCKFKKFKC 9  
 Db 36 RCLRLKIKC 44  
 :|: || ||  
 RESULT 5  
 SIP4\_YEAST  
 ID SIP4\_YEAST STANDARD; PRT; 829 AA.  
 AC P46954;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE SIP4 PROTEIN.  
 OS SIP4 OR YJL089W OR J0922.  
 GN SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMITASCOMYCETES; SACCHAROMYCETALES;  
 CC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RA LESAGE P., YANG X., CARLSON M.;

RL SUBMITTED (NOV-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-S288C;  
 RX MEDLINE; 96093911.  
 RA MIOGA T., SCHAEFF-GERSTENSCHLAEGER I., CHALWATZIS N., BAUR A.,  
 BOLES E., FOURNIER C., SCHMITT S., VELTEN C., WILHELM N.,  
 RA ZIMMERMANN F.K.;  
 RT "Sequence analysis of a 33.1 kb fragment from the left arm of  
 Saccharomyces cerevisiae chromosome X, including putative proteins  
 with leucine zippers, a fungal Zn(II)-Cys6 binuclear cluster domain  
 and a putative alpha 2-SCB-alpha 2 binding site.";  
 RL YEAST 11:681-689(1995).  
 CC -!- FUNCTION: INTERACTS WITH THE SNF1 PROTEIN KINASE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR  
 CLUSTER DOMAIN.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 -----  
 CC EMBL; U17643; G596255; -  
 DR EMBL; X83502; G929865; -  
 DR EMBL; Z49364; G1008258; -  
 DR SGD; L0001893; SIP4.  
 DR PROSITE; PS00463; ZN2\_CYS6\_FUNGAL\_1; 1.  
 DR PROSITE; PS00463; ZN2\_CYS6\_FUNGAL\_2; 1.  
 DR PFAM; PF00172; Zn.clus; 1.  
 DR HSP; P04386; 1AW6.  
 DR TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN; ZINC;  
 KW METAL-BINDING.  
 FT DNA\_BIND 46 73 ZN(2)-CYS(6), FUNGAL-TYPE.  
 FT CONFLICT 656 N -> T (IN REF. 1).  
 SQ SEQUENCE 829 AA; 96004 MW; 73D9AF3A CRC32;  
 -----  
 Query Match 61.7%; Score 37; DB 1; Length 829;  
 Best Local Similarity 55.6%; Pred. No. 40;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KCKFKKFKC 9  
 Db 48 RCLRLKIKC 56  
 :|: || ||  
 RESULT 6  
 KPC2\_DROME  
 ID KPC2\_DROME STANDARD; PRT; 700 AA.  
 AC P13677;  
 DT 01-APR-1990 (REL. 14, CREATED)  
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE PROTEIN KINASE C, EYE ISOZYME (EC 2.7.1.-) (PKC) (DPKC53E(EY))  
 DE (PROTEIN INAC) (PHOTORECEPTOR-SPECIFIC PKC) (EYE-PKC).  
 OS PKC2 OR INAC.  
 GN DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 CC DROSOPHILIDAE; DROSOPHILA.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 89249302.  
 RA SCHAEFFER E., SMITH D., MARDON G., QUINN W., ZUKER C.;  
 RT "Isolation and characterization of two new Drosophila protein kinase  
 C genes, including one specifically expressed in photoreceptor  
 cells.";  
 RL CELL 57:403-412(1989).  
 CC -!- FUNCTION: THIS IS A CALCIUM-ACTIVATED, PHOSPHOLIPID-DEPENDENT,

OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE: 95009532.  
 RX WEBSTER P.J., SUEN J., MACDONALD P.M.;  
 RT "Drosophila virilis oskar transgenes direct body patterning but not  
 RT pole cell formation or maintenance of mRNA localization in D.  
 RT melanogaster".  
 RL DEVELOPMENT 120:2027-2037(1994).  
 CC -1- FUNCTION: ORGANIZES THE GERM PLASM AND DIRECTS LOCALIZATION OF THE  
 CC POSTERIOR DETERMINANT NANOS. OSKAR PROTEIN IS REQUIRED TO KEEP  
 CC OSKAR RNA AND STAUFEN PROTEIN AT THE POSTERIOR POLE (BY  
 CC SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: POSTERIOR POLE OF THE OOCYTE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L22556; G520804; -.  
 DR FLYBASE; FBgn0013134; Dvir\osk.  
 KW DEVELOPMENTAL PROTEIN.  
 FT DOMAIN 42 46 POLY-GLN.  
 FT DOMAIN 49 59 POLY-GLN.  
 SQ SEQUENCE 594 AA; 68823 MW; FDB0BDD8 CRC32;  
 Query Match 65.0%; Score 39; DB 1; Length 594;  
 Best Local Similarity 75.0%; Pred. No. 15;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KCKFKFKC 8  
 Db 427 KCKFKFKC 434  
 RESULT 3  
 ID KMHA\_DICDI STANDARD; PRT: 1146 AA.  
 AC P42527;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE MYOSIN HEAVY CHAIN KINASE A (EC 2.7.1.129) (MHCK A).  
 GN MHKA OR MHCKA.  
 OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).  
 OC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=AX3;  
 RX MEDLINE: 95122486.  
 RA FUTEY L.M., MEDLEY Q.G., COTE G.P., EGELHOFF T.T.;  
 RT "Structural analysis of myosin heavy chain kinase A from  
 RT Dictyostelium. Evidence for a highly divergent protein kinase domain,  
 RT an amino-terminal coiled-coil domain, and a domain homologous to the  
 RT beta-subunit of heterotrimeric G proteins.";  
 RL J. BIOL. CHEM. 270:523-529(1995).  
 RN [2]  
 RP CHARACTERIZATION OF THE CATALYTIC DOMAIN.  
 RC STRAIN=AX3;  
 RX MEDLINE: 97207233.  
 RA COTE G.P., LUO X., MURPHY M.B., EGELHOFF T.T.;  
 RT "Mapping of the novel protein kinase catalytic domain of  
 RT Dictyostelium myosin II heavy chain kinase A";  
 RL J. BIOL. CHEM. 272:6846-6849(1997)  
 CC -1- FUNCTION: PHOSPHORYLATES THREONINE IN THE C-TERMINAL TAIL REGION  
 CC OF MYOSIN II HEAVY CHAIN. THIS PHOSPHORYLATION IS CRITICAL IN  
 CC REGULATING THE ASSEMBLY AND DISASSEMBLY OF MYOSIN II FILAMENT.

CC REQUIRES AUTOPHOSPHORYLATION FOR ACTIVITY.  
 CC -1- CATALYTIC ACTIVITY: ATP + [MYOSIN HEAVY-CHAIN] - ADP +  
 CC [MYOSIN HEAVY-CHAIN] PHOSPHATE.  
 CC -1- SUBUNIT: OLIGOMER.  
 CC -1- PTM: THE N-TERMINAL IS BLOCKED.  
 CC -1- COFACTOR: MN(2+) OR MG(2+) DEPENDENT.  
 CC -1- DOMAIN: CONSISTS OF AN N-TERMINAL DOMAIN WITH PROBABLE COILED COIL  
 CC STRUCTURE, A CENTRAL NONREPETITIVE CATALYTIC DOMAIN, AND A C-  
 CC TERMINAL DOMAIN WITH SEVEN WD REPEATS.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE MHCK / EF-2 PROTEIN KINASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U16856; G608520; -.  
 DR DICTYDB; DD01086; MHKA.  
 DR PROSITE; PS00678; WD-REPEATS; 4.  
 DR PFAM; PF00400; G-beta; 7.  
 KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING; REPEAT;  
 WD REPEAT; PHOSPHORYLATION; COILED COIL.  
 FT DOMAIN 100 120 COILED COIL (POTENTIAL).  
 FT DOMAIN 144 148 POLY-GLN.  
 FT DOMAIN 175 181 POLY-GLY.  
 FT DOMAIN 187 241 COILED COIL (POTENTIAL).  
 FT DOMAIN 297 502 COILED COIL (POTENTIAL).  
 FT DOMAIN 345 348 POLY-SER.  
 FT DOMAIN 438 441 POLY-LEU.  
 FT DOMAIN 500 551 PSEUDOSUBSTRATE/AUTOINHIBITORY DOMAIN  
 FT (POTENTIAL).  
 FT CATALYTIC.  
 FT NP\_BIND 552 852 ATP (POTENTIAL).  
 FT REPEAT 778 783 WD1.  
 FT REPEAT 867 897 WD2.  
 FT REPEAT 910 938 WD2.  
 FT REPEAT 952 980 WD3.  
 FT REPEAT 993 1021 WD4.  
 FT REPEAT 1033 1061 WD5.  
 FT REPEAT 1073 1101 WD6.  
 FT REPEAT 1114 1142 WD7.  
 SQ SEQUENCE 1146 AA; 128945 MW; A30841AA CRC32;  
 Query Match 63.3%; Score 38; DB 1; Length 1146;  
 Best Local Similarity 66.7%; Pred. No. 36;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KCKFKFKC 9  
 Db 47 QCPFKFKC 55  
 RESULT 4  
 ID PPRL\_YEAST STANDARD; PRT: 904 AA.  
 AC P07272;  
 DT 01-APR-1988 (REL. 07, CREATED)  
 DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE PYRIMIDINE PATHWAY REGULATORY PROTEIN 1.  
 GN PPRL OR YLR014C.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 85083082.  
 RA KAMMERER B., GUYONVARCH A., HUBERT J.-C.;  
 RT "Yeast regulatory gene PPRL1. I. Nucleotide sequence, restriction map

GenCore version 4.5

Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:20 ; Search time 71.87 Seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-124-280A-18

Perfect score: 60

Sequence: 1 KCKFKFKCK 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	66.7	299	1 SUT1_YEAST	P53032 saccharomyc
2	39	65.0	594	1 OSKA_DROVI	Q24741 drosophila
3	38	63.3	1146	1 KMHA_DICDI	P42527 dictyosteli
4	37	61.7	904	1 PPR1_YEAST	P07272 dictyosteli
5	37	61.7	829	1 SIP4_YEAST	P46954 saccharomyc
6	36	60.0	700	1 KPC2_DROME	P13677 drosophila
7	36	60.0	4543	1 LRP1_CHICK	P98157 gallus gall
8	36	60.0	852	1 RA54_SCHPO	P41410 schizosacch
9	36	60.0	1462	1 YAO6_CAEEL	Q20762 caenorhabdi
10	36	60.0	1459	1 YF1M_CAEEL	Q21874 caenorhabdi
11	35	58.3	320	1 AA3R_RAT	P28647 rattus norv
12	35	58.3	881	1 GAL4_YEAST	P04386 saccharomyc
13	35	58.3	1011	1 PPOL_CHICK	P26446 gallus gall
14	35	58.3	501	1 SUC1_CANAL	P33181 candida alb
15	35	58.3	712	1 SYTC_HUMAN	P26639 homo sapien
16	35	58.3	450	1 TH12_YEAST	P38141 saccharomyc
17	35	58.3	32	1 TX29_PHONI	P29426 phoneutria
18	35	58.3	93	1 VIA_TAV	P28931 tomato aspe
19	35	58.3	637	1 YAO7_SCHPO	Q10086 schizosacch
20	34.5	57.5	4753	1 LRP_CAEEL	Q04833 caenorhabdi
21	34	56.7	970	1 DAB1_YEAST	P21657 saccharomyc
22	34	56.7	863	1 ENV_RV128	P05882 human immun
23	34	56.7	524	1 HUNE_TRICA	Q01791 tribolium c
24	34	56.7	61	1 MTH_HUMAN	P80294 homo sapien
25	34	56.7	184	1 RAPA_DISOM	P22123 discopyge o
26	34	56.7	259	1 RM02_PARTI	P15765 paramecium
27	34	56.7	582	1 TATR_NPVAC	P11138 autographa
28	34	56.7	587	1 TATR_NPVAC	P33245 bombyx mori
29	34	56.7	880	1 YJF6_YEAST	P47043 saccharomyc
30	33	55.0	252	1 L433_DICDI	P54632 dictyosteli
31	33	55.0	254	1 ADH_DROSU	Q03384 drosophila
32	33	55.0	363	1 ALF_SCHMA	P53442 schistosoma
33	33	55.0	865	1 LAC9_KLULA	P08657 kluyveromyc
34	33	55.0	1456	1 MANR_HUMAN	P22897 homo sapien
35	33	55.0	265	1 MPH1_HOLLA	P43216 holcus lana
36	33	55.0	750	1 PTP2_YEAST	P29461 saccharomyc
37	33	55.0	1170	1 YK08_YEAST	P32862 saccharomyc
38	33	55.0	416	1 YMW7_YEAST	P54730 saccharomyc
39	32	53.3	499	1 ZABA_DROME	P36872 drosophila
40	32	53.3	447	1 ZABA_DROME	Q00007 homo sapien
41	32	53.3	441	1 ZABA_HUMAN	P36876 rattus norv
42	32	53.3	443	1 ZABA_RAT	Q00005 homo sapien
43	32	53.3	443	1 ZABB_HUMAN	P54614 sus scrofa

44 32 53.3 413 1 ZABB\_RABIT Q00006 oryctolagus  
45 32 53.3 1060 1 ZFH1\_DROME P28166 drosophila

## ALIGNMENTS

## RESULT 1

SUT1\_YEAST ID SUT1\_YEAST STANDARD; PRT; 299 AA.  
AC P53032;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PROBABLE STEROL CARRIER  
GN SUT1 OR YGL162W OR G1828.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
ON SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 28383 / FL100;  
RX MEDLINE; 96084960.  
RA BOURROT S., KARST F.;  
RT "Isolation and characterization of the Saccharomyces cerevisiae SUT1  
RT gene involved in sterol uptake.";  
RL GENE 165:97-102(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96158061.  
RA JAMES C.M., INDGE K.J., OLIVER S.G.;  
RT "DNA sequence analysis of a 35 kb segment from Saccharomyces  
RT cerevisiae chromosome VII reveals 19 open reading frames including  
RT RAD54, ACE1/CUP2, PMK1, RCK1, AMS1 and CAL1/CDC43.";  
RL YEAST 11:1413-1419(1995).  
CC -!- FUNCTION: INVOLVED IN STEROL UPTAKE.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; X77766; E99231;  
DR EMBL; 248618; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; 272684; E243599; -  
DR SGD; L0003483; SUT1.  
KW TRANSPORT.  
SQ SEQUENCE 299 AA; 33191 MW; D0F553A8 CRC32;

Query Match 66.7%; Score 40; DB 1; Length 299;  
Best Local Similarity 66.7%; Pred. No. 5.7;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KCKFKFKC 9

||: || ||  
Db 162 KCLRLKIKC 170

## RESULT 2

OSKA\_DROVI ID OSKA\_DROVI STANDARD; PRT; 594 AA.  
AC Q24741;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MATERNAL EFFECT PROTEIN OSKAR.  
GN OSK.  
OS DROSOPHILA VIRILIS (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;

C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 12-Dec-1997  
 C:Accession: A05022; S61016; S65277; S12977  
 R:Laughon, A.; Gesteland, R.F.  
 Mol. Cell. Biol. 4, 260-267, 1984  
 A:Title: Primary structure of the *Saccharomyces cerevisiae* GAL4 gene.  
 A:Reference number: A05022; MUID:84141879  
 A:Accession: A05022  
 A:Molecule type: DNA  
 A:Residues: 1-881 <LAU>  
 A:Cross-references: EMBL:K01486; NID:g171557; PID:g171558  
 R:Gadhavi, P.L.; Raine, A.R.C.; Alefounder, P.R.; Laue, E.D.  
 FEBS Lett. 276, 49-53, 1990  
 A:Title: Complete assignment of the <sup>1</sup>H NMR spectrum and secondary structure of the DNA  
 A:Reference number: S12977; MUID:91092433  
 A:Contents: annotation; zinc finger  
 R:Pohl, T.M.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: S61010  
 A:Accession: S61016  
 A:Molecule type: DNA  
 A:Residues: 1-881 <POH>  
 A:Cross-references: EMBL:267751; NID:g1061234; PID:g1061241  
 R:Pohl, T.M.  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S64899  
 A:Accession: S65277  
 A:Molecule type: DNA  
 A:Residues: 1-881 <POW>  
 A:Cross-references: EMBL:273604; NID:g1370510; PID:e246953; PID:g1370511; MIPS:YPL248C  
 A:Experimental source: strain S288C (A8972)  
 C:Comment: This protein is a positive regulator for the gene expression of the galactose  
 C:Genetics:  
 A:Gene: SGD:GAL4  
 A:Cross-references: SGD:S0006169; MIPS:YPL248C  
 A:Map position: 16L  
 C:Superfamily: regulatory protein GAL4; GAL4 zinc binuclear cluster homology  
 C:Keywords: DNA binding; galactose utilization; transcription regulation; zinc finger  
 F:6-43/Domain: GAL4 zinc binuclear cluster homology <GAL4>  
 F:11-38/Region: zinc finger CCCC motif

Query Match 58.3%; Score 35; DB 1; Length 881;  
 Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 CKFKFKC 9  
 I: I I I  
 Db 14 CRLKLLKC 21

Search completed: September 7, 1999, 23:06:32  
 Job time: 2480 sec

```

Matches      6;  Conservative      1;  Mismatches      3;  Indels      0;  Gaps      0;

QY      1  KCKFKFKCK 10
      || ||| | :
Db      16  KCSFKKLKAE 25

RESULT 11
B69325
LPS biosynthesis protein homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Jun-1998
C:Accession: B69325
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: B69325
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <KLE>
A:Cross-references: GB:AE001063; GB:AE000782; NID:g2689386; PID:g2650016; TIGR:AF0602

Query Match      60.0%;  Score 36;  DB 2;  Length 396;
Best Local Similarity 66.7%;  Pred. No. 57;
Matches      8;  Conservative      1;  Mismatches      1;  Indels      2;  Gaps      1;

QY      1  KCKFKFKCK-CK 10
      | |||: || ||
Db      83  KAKFKQFKDKCK 94

RESULT 12
S41886
DNA repair protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 09-Sep-1997
C:Accession: S41886; S41444
R:Muris, D.F.R.; Vreeken, K.; Smit, C.; Carr, A.M.; Broughton, B.C.; Lehmann, A.R.; Lohm
submitted to the EMBL Data Library, January 1994
A:Description: Isolation of the Schizosaccharomyces pombe RAD54 homologue, a gene involv
A:Reference number: S41444
A:Accession: S41886
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852 <MURL>
A:Cross-references: EMBL:229640; NID:g452689; PID:g452690
R:Muris, D.F.R.; Vreeken, K.; Smit, C.; Carr, A.M.; Broughton, B.C.; Lehmann, A.R.; Lohm
submitted to the EMBL Data Library, January 1994
A:Description: Isolation of the Schizosaccharomyces pombe RAD54 homologue, a gene involv
A:Reference number: S41444
A:Accession: S41444
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852 <MUR2>
A:Cross-references: EMBL:229615
C:Keywords: DNA binding; nucleus

Query Match      60.0%;  Score 36;  DB 2;  Length 852;
Best Local Similarity 100.0%;  Pred. No. 1e+02;
Matches      6;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      4  FKFKCK 9
      ||||| |
Db      34  FKFKCK 39

Matches      6;  Conservative      1;  Mismatches      3;  Indels      0;  Gaps      0;

QY      1  KCKFKFKCK 10
      || ||| | :
Db      16  KCSFKKLKAE 25

RESULT 11
B69325
LPS biosynthesis protein homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Jun-1998
C:Accession: B69325
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: B69325
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <KLE>
A:Cross-references: GB:AE001063; GB:AE000782; NID:g2689386; PID:g2650016; TIGR:AF0602

Query Match      60.0%;  Score 36;  DB 2;  Length 396;
Best Local Similarity 66.7%;  Pred. No. 57;
Matches      8;  Conservative      1;  Mismatches      1;  Indels      2;  Gaps      1;

QY      1  KCKFKFKCK-CK 10
      | |||: || ||
Db      83  KAKFKQFKDKCK 94

RESULT 12
S41886
DNA repair protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 09-Sep-1997
C:Accession: S41886; S41444
R:Muris, D.F.R.; Vreeken, K.; Smit, C.; Carr, A.M.; Broughton, B.C.; Lehmann, A.R.; Lohm
submitted to the EMBL Data Library, January 1994
A:Description: Isolation of the Schizosaccharomyces pombe RAD54 homologue, a gene involv
A:Reference number: S41444
A:Accession: S41886
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852 <MURL>
A:Cross-references: EMBL:229640; NID:g452689; PID:g452690
R:Muris, D.F.R.; Vreeken, K.; Smit, C.; Carr, A.M.; Broughton, B.C.; Lehmann, A.R.; Lohm
submitted to the EMBL Data Library, January 1994
A:Description: Isolation of the Schizosaccharomyces pombe RAD54 homologue, a gene involv
A:Reference number: S41444
A:Accession: S41444
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852 <MUR2>
A:Cross-references: EMBL:229615
C:Keywords: DNA binding; nucleus

Query Match      60.0%;  Score 36;  DB 2;  Length 852;
Best Local Similarity 100.0%;  Pred. No. 1e+02;
Matches      6;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      4  FKFKCK 9
      ||||| |
Db      34  FKFKCK 39

```

```

RESULT 13
T00390
KIAA0614 protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 01-Feb-1999
C:Accession: T00390
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura,
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The compl
A:Reference number: Z14142
A:Accession: T00390
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1630 <ISH>
A:Cross-references: EMBL:AB014514; NID:d1204269; PID:d1032550
C:Genetics:
A:Note: KIAA0614

Query Match      60.0%;  Score 36;  DB 2;  Length 1630;
Best Local Similarity 75.0%;  Pred. No. 1.7e+02;
Matches      6;  Conservative      0;  Mismatches      2;  Indels      0;  Gaps      0;

QY      2  CKFKFKCK 9
      ||| || |
Db      1543  CKFKFKAC 1550

```

```

RESULT 14
YSHUT
threonine--trna ligase (EC 6.1.1.3) - human
N:Alternate names: threonyl-trna synthetase
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1992 #sequence_revision 30-Jun-1993 #text_change 05-Sep-1997
C:Accession: A38867; A40017
R:Cruzen, M.E.; Arfin, S.M.
submitted to the EMBL Data Library, June 1991
A:Reference number: A38867
A:Accession: A38867
A:Molecule type: mRNA
A:Residues: 1-712 <CRU>
A:Cross-references: EMBL:M63180; NID:g339679; PID:g1464742
R:Cruzen, M.E.; Arfin, S.M.
J. Biol. Chem. 266, 9919-9923, 1991
A:Title: Nucleotide and deduced amino acid sequence of human threonyl-trna synthetase
A:Reference number: A40017; MUID:91236775
A:Accession: A40017
A:Molecule type: mRNA
A:Residues: 1-66, 'EC', '69-274', 'A', '276-385', 'F', '387-454', 'F', '456-506', 'R', '508-527', 'C', '529-
A:Cross-references: GB:M63180
A:Note: the authors translated the codon GGC for residue 152 as Ala
C:Genetics:
A:Gene: GDB:TARS
A:Cross-references: GDB:l120397; OMIM:187790
A:Map position: 5p13-5cen
C:Superfamily: threonine--trna ligase
C:Keywords: aminoacyl-trna synthetase; ATP; ligase; protein biosynthesis

Query Match      58.3%;  Score 35;  DB 1;  Length 712;
Best Local Similarity 62.5%;  Pred. No. 1.3e+02;
Matches      5;  Conservative      2;  Mismatches      1;  Indels      0;  Gaps      0;

QY      3  KFKFKCK 10
      | : |||| |
Db      219  KYNFKCKR 226

```

```

RESULT 15
RGBYG4
regulatory protein GAL4 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein Pl021; protein YPL248C
C:Species: Saccharomyces cerevisiae

Query Match      58.3%;  Score 35;  DB 1;  Length 712;
Best Local Similarity 62.5%;  Pred. No. 1.3e+02;
Matches      5;  Conservative      2;  Mismatches      1;  Indels      0;  Gaps      0;

QY      3  KFKFKCK 10
      | : |||| |
Db      219  KYNFKCKR 226

```



Db | ::::|::|  
16 KPRIYKCK 25

RESULT 9  
A32392  
protein kinase C (EC 2.7.1.1-) 535(ey) - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C:date: 22-Nov-1989 #sequence\_revision 22-Nov-1989 #text\_change 24-Sep-1998  
C:Accession: A32392  
R:Schaeffer, E.; Smith, D.; Mardon, G.; Quinn, W.; Zuker, C.  
Cell 57, 403-412, 1989  
A:title: Isolation and characterization of two new *Drosophila* protein kinase C genes,  
A:reference number: A32392; MUID:89249302  
A:Accession: A32392  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-700 <SCH>  
A:Cross-references: GB:J04845; NID:g158126; PID:g158127  
C:Genetics:  
A:Gene: FlyBase:InAc  
A:Cross-references: FlyBase:FBN0004784  
C:Superfamily: protein kinase C alpha; protein kinase C C2 region homology; protein k  
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F:72-121/Domain: protein kinase C zinc-binding repeat homology <K21>  
F:137-186/Domain: protein kinase C zinc-binding repeat homology <K22>  
F:187-299/Domain: protein kinase C C2 region homology <KC2>  
F:369-629/Domain: protein kinase homology <KIN>  
F:377-385/Region: protein kinase ATP-binding motif

```

Query Match          60.0%; Score 36; DB 2; Length 700;
Best Local Similarity 75.0%;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      2 CKFKKFKC 9
         ||| |||
Db      114 CKFVVFKC 121

```

```

RESULT 10
SG2853
hypothetical protein D02_orf122b - Mycoplasma pneumoniae (SGC3)
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 18-Sep-1998
C:Accession: SG2853; S73575
R:Hilbert, H.; Himmelreich, R.; Plagens, H.; Herrmann, R.
Nucleic Acids Res. 24, 628-639, 1996
A:Title: Sequence analysis of 56 kb from the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S63797; MUID:96177562
A:Accession: S62853
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-122 <HIL>
A:Cross-references: EMBL:U43738; NID:g1209757; PID:g1209770
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S73327; MUID:97105885
A:Accession: S73575
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-122 <HIM>
A:Cross-references: EMBL:AE000024; GB:U00089; NID:g1673904; PID:g1673917
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Genetic code: SGC3

```

Query Match	60.0%;	Score 36;	DB 2;	Length 122;
Best Local Similarity	60.0%;	Pred. No. 22;		

A:Residues: 1-655,'N',657-829 <MIW>  
A:Cross-references: EMBL:249364; NID:g1008257; PID:g1008258; MIPS:YJL089w  
C:Genetics:  
A:Gene: SGD:SIP4  
A:Cross-references: SGD:S0003625; MIPS:YJL089w  
A:Map position: 10L  
C:Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluster h  
F:41-78/Domain: GAL4 zinc binuclear cluster homology <GAL4>  
  
Query Match 61.7%; Score 37; DB 2; Length 829;  
Best Local Similarity 55.6%; Pred. No. 71;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KCKFKKFKC 9  
Db 48 RCRLKKIKC 56  
  
RESULT 6  
F64393  
hypothetical protein MJ0750 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 10-Oct-1997  
C:Accession: F64393  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
Rson, J.D.; Sadow, P.W.; Hannu, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999  
A:Accession: F64393  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-238 <BUL>  
A:Cross-references: GB:U67521; GB:L77117; NID:g1591463; PID:g1592304; TIGR:MU0750; PID:9  
C:Genetics:  
A:Map position: REV676712-675996  
  
Query Match 61.7%; Score 37; DB 2; Length 238;  
Best Local Similarity 66.7%; Pred No. 26;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KCKFKKFKC 9  
Db 190 KCKLCEFKC 198  
  
RESULT 7  
A53102  
alpha-2-macroglobulin receptor precursor - chicken  
N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein recepto  
C:Species: Gallus gallus (chicken)  
C:Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 05-Dec-1998  
C:Accession: A53102  
R:Nimf, J.; Stifani, S.; Bilous, P.T.; Schneider, W.J.  
J. Biol. Chem. 269, 212-219, 1994  
A:Title: The somatic cell-specific low density lipoprotein receptor-related protein of t  
A:Reference number: A53102; MUID:94103212  
A:Accession: A53102  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-4543 <NTM>  
A:Cross-references: GB:X74904; NID:g438006; PID:g438007  
C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associat  
d protein.  
C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding  
C:Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycopro  
F:11-17/Domain: signal sequence #status predicted <SIG>  
F:18-3942/Domain: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>  
F:18-3942,3943-4543/Product: alpha-2-macroglobulin receptor #status predicted <MAT>

F:29-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:74-110/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:117-150/Domain: EGF homology <EG1>  
F:156-190/Domain: EGF homology <EG2>  
F:200-241/Domain: LDL receptor WYTD-containing repeat homology <YW01>  
F:243-283/Domain: LDL receptor WYTD-containing repeat homology <YW02>  
F:294-336/Domain: LDL receptor WYTD-containing repeat homology <YW03>  
F:337-380/Domain: LDL receptor WYTD-containing repeat homology <YW04>  
F:381-422/Domain: LDL receptor WYTD-containing repeat homology <YW05>  
F:423-470/Domain: LDL receptor WYTD-containing repeat homology <YW06>  
F:480-521/Domain: EGF homology <EG3>  
F:573-615/Domain: LDL receptor WYTD-containing repeat homology <YW07>  
F:616-661/Domain: LDL receptor WYTD-containing repeat homology <YW08>  
F:662-712/Domain: LDL receptor WYTD-containing repeat homology <YW09>  
F:713-754/Domain: LDL receptor WYTD-containing repeat homology <YW10>  
F:755-797/Domain: LDL receptor WYTD-containing repeat homology <YW11>  
F:805-840/Domain: EGF homology <EG4>  
F:852-888/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F:893-929/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F:934-969/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F:974-1009/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F:1013-1049/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F:1060-1095/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
F:1102-1138/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
F:1143-1180/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
F:1183-1219/Domain: EGF homology <EG5>  
F:1225-1259/Domain: EGF homology <EG6>  
F:1267-1306/Domain: LDL receptor WYTD-containing repeat homology <YW12>  
F:1307-1353/Domain: LDL receptor WYTD-containing repeat homology <YW13>  
F:1354-1396/Domain: LDL receptor WYTD-containing repeat homology <YW14>  
F:1397-1443/Domain: LDL receptor WYTD-containing repeat homology <YW15>  
F:1444-1486/Domain: LDL receptor WYTD-containing repeat homology <YW16>  
F:1487-1529/Domain: LDL receptor WYTD-containing repeat homology <YW17>  
F:1538-1576/Domain: EGF homology <EG7>  
F:1591-1624/Domain: LDL receptor WYTD-containing repeat homology <YW18>  
F:1625-1667/Domain: LDL receptor WYTD-containing repeat homology <YW19>  
F:1668-1711/Domain: LDL receptor WYTD-containing repeat homology <YW20>  
F:1712-1751/Domain: LDL receptor WYTD-containing repeat homology <YW21>  
F:1752-1794/Domain: LDL receptor WYTD-containing repeat homology <YW22>  
F:1795-1842/Domain: LDL receptor WYTD-containing repeat homology <YW23>  
F:1846-1882/Domain: EGF homology <EG8>  
F:1930-1972/Domain: LDL receptor WYTD-containing repeat homology <YW24>  
F:1973-2015/Domain: LDL receptor WYTD-containing repeat homology <YW25>  
F:2016-2059/Domain: LDL receptor WYTD-containing repeat homology <YW26>  
F:2060-2101/Domain: LDL receptor WYTD-containing repeat homology <YW27>  
F:2102-2147/Domain: LDL receptor WYTD-containing repeat homology <YW28>  
F:2155-2190/Domain: EGF homology <EG9>  
F:2195-2237/Domain: LDL receptor WYTD-containing repeat homology <YW29>  
F:2247-2288/Domain: LDL receptor WYTD-containing repeat homology <YW30>  
F:2338-2382/Domain: LDL receptor WYTD-containing repeat homology <YW31>  
F:2383-2423/Domain: LDL receptor WYTD-containing repeat homology <YW32>  
F:2424-2467/Domain: LDL receptor WYTD-containing repeat homology <YW33>  
F:2476-2511/Domain: EGF homology <EG10>  
F:2518-2555/Domain: LDL receptor ligand-binding repeat homology <LDL1B>  
F:2560-2594/Domain: LDL receptor ligand-binding repeat homology <LDL1C>  
F:2599-2633/Domain: LDL receptor ligand-binding repeat homology <LDL1D>  
F:2646-2682/Domain: LDL receptor ligand-binding repeat homology <LDL1E>  
F:2690-2724/Domain: LDL receptor ligand-binding repeat homology <LDL1F>  
F:2732-2767/Domain: LDL receptor ligand-binding repeat homology <LDL1G>  
F:2772-2810/Domain: LDL receptor ligand-binding repeat homology <LDL1H>  
F:2816-2851/Domain: LDL receptor ligand-binding repeat homology <LDL1I>  
F:2856-2895/Domain: LDL receptor ligand-binding repeat homology <LDL1J>  
F:2902-2936/Domain: LDL receptor ligand-binding repeat homology <LDL1K>  
F:2941-2977/Domain: EGF homology <EG11>  
F:2983-3018/Domain: EGF homology <EG12>  
F:3026-3065/Domain: LDL receptor WYTD-containing repeat homology <YW34>  
F:3066-3110/Domain: LDL receptor WYTD-containing repeat homology <YW35>  
F:3111-3153/Domain: LDL receptor WYTD-containing repeat homology <YW36>  
F:3154-3197/Domain: LDL receptor WYTD-containing repeat homology <YW37>  
F:3198-3238/Domain: LDL receptor WYTD-containing repeat homology <YW38>  
F:3239-3281/Domain: LDL receptor WYTD-containing repeat homology <YW39>  
F:3291-3327/Domain: EGF homology <EG13>  
F:3331-3366/Domain: LDL receptor ligand-binding repeat homology <LDL1>

S.  
 A:Reference number: S74322; MUID:97061201  
 A:Accession: S76182  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-243 <KAN>  
 A:Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PID:d1019174; PID:g1653528  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 63.3%; Score 38; DB 2; Length 243;  
 Best Local Similarity 85.7%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 KKKFKFC 9  
 |||||  
 DB 88 KKKFKFC 94

RESULT 3  
 A5532  
 myosin-heavy-chain kinase (EC 2.7.1.129) A - slime mold (Dictyostelium discoideum)  
 C:Species: Dictyostelium discoideum  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Apr-1998  
 C:Accession: A5532  
 R:Futey, L.M.; Medley, Q.G.; Cote, G.P.; Egelhoff, T.T.  
 J. Biol. Chem. 270, 523-529, 1995  
 A:Title: Structural analysis of myosin heavy chain kinase A from Dictyostelium. Evidence  
 the beta-subunit of heterotrimeric G proteins.  
 A:Reference number: A5532; MUID:95122486  
 A:Accession: A5532  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1146 <FUT>  
 A:Cross-references: GB:U16856; NID:g608519; PID:g608520  
 C:Genetics:  
 A:introns: #status absent  
 C:Superfamily: WD repeat homology  
 C:Keywords: autophosphorylation; coiled coil; multimer; phosphoprotein; phosphotransferase  
 F:865-898/Domain: WD repeat homology <WD1>  
 F:950-981/Domain: WD repeat homology <WD2>  
 F:991-1022/Domain: WD repeat homology <WD3>  
 F:1031-1062/Domain: WD repeat homology <WD4>  
 F:1071-1102/Domain: WD repeat homology <WD5>  
 F:1112-1143/Domain: WD repeat homology <WD6>

Query Match 63.3%; Score 38; DB 2; Length 1146;  
 Best Local Similarity 86.7%; Pred. No. 64;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KCKFKFKFC 9  
 :|||  
 DB 47 QCPFKFKFC 55

RESULT 4  
 RGBYP1  
 regulatory protein PPR1 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein L1575; protein YLR014c; pyrimidine pathway regulatory protein  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 26-Feb-1999  
 C:Accession: S05877; S59276; S59799; S64836  
 R:Kammerer, B.; Guyonvarch, A.; Hubert, J.C.  
 J. Mol. Biol. 180, 239-250, 1994  
 A:Title: Yeast regulatory gene PPR1. I. Nucleotide sequence, restriction map and codon u  
 A:Reference number: S05877; MUID:85083082  
 A:Accession: S05877  
 A:Molecule type: DNA  
 A:Residues: 1-904 <KAN>  
 A:Cross-references: EMBL:X01739; NID:g4214; PID:g4215  
 R:Saville, S.P.; Atkinson, S.; Jamieson, L.; Pocklington, M.J.; Orr, E.  
 submitted to the EMBL Data Library, August 1995

A:Description: A 7.8kb fragment from chromosome XII of Saccharomyces cerevisiae does  
 A:Reference number: S59270  
 A:Accession: S59276  
 A:Molecule type: DNA  
 A:Residues: 1-513 <SAV>  
 A:Cross-references: EMBL:X90564; NID:g975221; PID:g975228  
 A:Experimental source: strain S288C  
 R:Liljelund, P.; Losson, R.; Kammerer, B.; Lacroute, F.  
 J. Mol. Biol. 180, 251-265, 1984  
 A:Title: Yeast regulatory gene PPR1. II. Chromosomal localization, meiotic map, suppr  
 A:Reference number: S59799; MUID:85083083  
 A:Accession: S59799  
 A:Molecule type: DNA  
 A:Residues: 1-75 <LIL>  
 A:Cross-references: EMBL:M29131; NID:g342049; PID:g903888  
 R:Vandenbol, M.; Portetelle, D.; Hilger, F.  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S64742  
 A:Accession: S64836  
 A:Molecule type: DNA  
 A:Residues: 1-904 <VAN>  
 A:Cross-references: EMBL:Z73186; NID:g1360311; PID:e245770; PID:g1360312; MIPS:YLR014  
 A:Note: experimental\_source strain S288C  
 C:Genetics:  
 A:Gene: SGD:PPR1  
 A:Cross-references: SGD:S0004004; MIPS:YLR014c  
 A:Map position: 12R  
 C:Superfamily: regulatory protein PPR1; GAL4 zinc binuclear cluster homology  
 C:Keywords: DNA binding; transcription regulation; zinc finger  
 F:29-66/Domain: GAL4 zinc binuclear cluster homology <GAL4>  
 F:34-61/Region: zinc finger CCCC motif

Query Match 61.7%; Score 37; DB 1; Length 904;  
 Best Local Similarity 55.6%; Pred. No. 76;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KCKFKFKFC 9  
 :|||  
 DB 36 RCRLKKIKC 44

RESULT 5  
 S50246  
 SIP4 protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein J0922; protein YJL089w  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 13-Jan-1995 #sequence\_revision 10-Feb-1995 #text\_change 17-Mar-1999  
 C:Accession: S50246; S56019; S56866  
 R:Lesage, P.; Yang, X.; Carlson, M.  
 submitted to the EMBL Data Library, November 1994  
 A:Description: SIP4, a protein that interacts with the SNF1 protein kinase.  
 A:Reference number: S50246  
 A:Accession: S50246  
 A:Molecule type: DNA  
 A:Residues: 1-829 <LES>  
 A:Cross-references: EMBL:U17643; NID:g596254; PID:g596255  
 R:Miosga, T.; Schaaff-Gerstenschlaeger, I.; Chaltatzis, N.; Baur, A.; Boles, E.; Four  
 Yeast 11, 681-689, 1995  
 A:Title: Sequence analysis of a 33.1 kb fragment from the left arm of Saccharomyces c  
 ter domain and a putative alpha-2-SCB-alpha-2 binding site.  
 A:Reference number: S56016; MUID:96093911  
 A:Accession: S56019  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-655, N', 657-829 <MIO>  
 A:Cross-references: EMBL:X83502; NID:g929861; PID:g929865  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994  
 R:Miosga, T.; Schaaff-Gerstenschlaeger, I.; Baur, A.; Boles, E.; Chaltatzis, N.; Four  
 submitted to the Protein Sequence Database, September 1995  
 A:Reference number: S56855  
 A:Accession: S56866  
 A:Molecule type: DNA

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:31 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-18  
Perfect score: 60  
Sequence: 1 KCKFKFKCK 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	66.7	299	JC4374	sterol uptake prot
2	38	63.3	243	S76182	hypothetical prote
3	38	63.3	1146	A55352	myosin-heavy-chain
4	37	61.7	904	1 RGBXP1	regulatory protein
5	37	61.7	829	S50246	SIP4 protein - yea
6	37	61.7	238	F64393	hypothetical prote
7	36	60.0	4543	1 A53102	alpha-2-macroglobu
8	36	60.0	80	G64391	rubredoxin - Metha
9	36	60.0	700	2 A32392	protein kinase c (
10	36	60.0	122	S62853	hypothetical prote
11	36	60.0	396	2 B69325	LPS biosynthesis p
12	36	60.0	852	2 S41886	DNA repair protein
13	36	60.0	1630	2 T00390	KIAA0614 protein -
14	35	58.3	712	1 YSHUT	threonine--crna li
15	35	58.3	881	1 RGBIG4	regulatory protein
16	35	58.3	993	1 PIVXTA	RNA i protein - to
17	35	58.3	1011	2 JH0581	NAD+ ADP-ribosyltr
18	35	58.3	506	2 A40218	envelop glycoprote
19	35	58.3	501	2 A43302	probable finger pr
20	35	58.3	450	2 S46117	probable regulator
21	35	58.3	337	2 S68678	adenosine receptor
22	35	58.3	320	2 S17177	probable G protein
23	35	58.3	228	2 H70250	hypothetical prote
24	35	58.3	268	2 S57542	hypothetical prote
25	35	58.3	32	2 S29217	neurotoxin Tx2 - s
26	35	58.3	562	2 S56145	BS69 protein - hum
27	34	57.5	4733	2 A47437	LDL-receptor-relat
28	34	56.7	259	1 B64453	probable pyruvate
29	34	56.7	229	1 R5PPL2	ribosomal protein
30	34	56.7	581	1 RGNVBV	trans-activating t
31	34	56.7	582	1 RGNVE2	GTP-binding protei
32	34	56.7	184	2 F38625	metallothionein IH
33	34	56.7	61	2 S47651	transregulatory pr
34	34	56.7	582	2 A49626	immediate early pr
35	34	56.7	587	2 S20596	env protein - huma
36	34	56.7	219	2 S25939	env protein - huma
37	34	56.7	211	2 S25938	probable membrane
38	34	56.7	769	2 S50966	transcription acti
39	34	56.7	970	2 S48485	

hypothetical prote  
finger protein YUL  
hypothetical prote  
phosphoprotein pho  
tumor necrosis fac  
terminal protein p

40 34 56.7 720 2 T02734  
41 34 56.7 880 2 S56828  
42 34 56.7 1398 2 H71606  
43 34 56.7 443 2 JS0752  
44 34 56.7 557 2 JC6539  
45 34 56.7 641 2 JC5648

ALIGNMENTS

RESULT 1  
JC4374  
sterol uptake protein 1 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: Protein G1828; protein YGL162w; SUT1 protein  
C:Species: Saccharomyces cerevisiae  
C>Date: 10-Jan-1996 #sequence\_revision 08-Feb-1996 #text\_change 06-Feb-1998  
C:Accession: JC4374; S60424; S64178  
R:Bourrot, S.; Karst, F.  
Gene 165, 97-102, 1995  
A:Title: Isolation and characterization of the Saccharomyces cerevisiae SUT1 gene inv  
A:Reference number: JC4374; MUID:96084960  
A:Accession: JC4374  
A:Molecule type: DNA  
A:Residues: 1-299 <BOU>  
A:Cross-references: EMBL:X77766; NID:g1183994; PID:e99231; PID:g1183995  
A:Experimental source: SBI  
R:James, C.M.; Indge, K.J.; Oliver, S.G.  
Yeast 11, 1413-1419, 1995  
A:Title: DNA sequence analysis of a 35 kb segment from Saccharomyces cerevisiae chrom  
A:Reference number: S60417; MUID:96158061  
A:Accession: S60424  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-299 <JAM>  
A:Cross-references: EMBL:Z48618  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995  
R:James, C.M.; Indge, K.J.; Oliver, S.G.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64165  
A:Accession: S64178  
A:Molecule type: DNA  
A:Residues: 1-299 <JAM>  
A:Cross-references: EMBL:Z72684; NID:g1322757; PID:e243599; PID:g1322758; MIPS:YGL162  
A:Experimental source: strain S388C  
C:Comment: this protein is a member of the hypoxic gene family, and is rich in serine  
C:Genetics:  
A:Gene: SGD:SUT1  
A:Cross-references: SGD:S0003130; MIPS:YGL162w  
A:Map position: 7L  
F:133-151/Region: serine/threonine-rich

Query Match 66.7%; Score 40; DB 2; Length 299;  
Best Local Similarity 66.7%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KCKFKFKCK 9  
||: ||: ||  
DB 162 KCRLKRAIKC 170

RESULT 2  
S76182  
hypothetical protein - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
C:Accession: S76182  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys

**THIS PAGE BLANK (USPT)**

```

: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/714,070A
: FILING DATE: 07-OCT-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94/01767
: FILING DATE: 10-MAR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Rea, Teresa Stanek
: REGISTRATION NUMBER: 30,427
: REFERENCE/DOCKET NUMBER: 017753-077
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 775 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Schizosaccharomyces pombe
: US-08-714-070A-1

```

```

Query Match          53.3%; Score 32; DB 2; Length 775;
Best Local Similarity 62.5%; Pred. No. 5.3e-02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      2 CKFKKFKC 9
       1: |||
Db      42 CROKKIKC 49

```

```

Search completed: September 7, 1999, 22:38:38
Job time: 7922 sec

```

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-175-096-2

Query Match 55.0%; Score 33; DB 1; Length 261;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KFKKFKCK 10  
| | : | | |  
Db 156 KFRVKCK 163

## RESULT 13

US-08-436-463-6  
; Sequence 6, Application US/08436463  
; Patent No. 5760185

## GENERAL INFORMATION:

APPLICANT: KIMACHI, Kazuhiko

APPLICANT: MAEDA, Hiroaki

APPLICANT: NISHIYAMA, Kiyoto

APPLICANT: TOKIYOSHI, Sachio

TITLE OF INVENTION: ANTI-FELINE HERPES VIRUS-1 RECOMBINANT

TITLE OF INVENTION: ANTIBODY AND GENE FRAGMENT CODING FOR SAID ANTIBODY

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK, P. L. L.C.

STREET: 419 Seventh Street, N.W., Suite 400

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/436,463

FILING DATE: 26-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 341255/1992

FILING DATE: 28-NOV-1992

ATTORNEY/AGENT INFORMATION:

NAME: YUN, Allen C.

REGISTRATION NUMBER: 37,971

REFERENCE/DOCKET NUMBER: KIMACHI-1

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 333 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-436-463-6

Query Match 55.0%; Score 33; DB 2; Length 333;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KFKKFKCK 10  
| | : | | |  
Db 199 KGKFKCK 206

## RESULT 14

US-08-024-253-6

; Sequence 6, Application US/08024253  
; Patent No. 5785968

## GENERAL INFORMATION:

APPLICANT: KIMACHI, Kazuhiko

APPLICANT: MAEDA, Hiroaki

APPLICANT: NISHIYAMA, Kiyoto

APPLICANT: TOKIYOSHI, Sachio

APPLICANT: TOHYA, Yukinobu

APPLICANT: MIKAMI, Takeshi

TITLE OF INVENTION: ANTI-FELINE CALCIVIRUS RECOMBINANT

TITLE OF INVENTION: ANTIBODY AND GENE FRAGMENT ENCODING THE SAME

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: WEGNER, CANTOR, MUELLER & PLAYER

STREET: 1233 20th Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20036-8218

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/024,253

FILING DATE: 19930301

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 79189/1992

FILING DATE: 28-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: CANTOR, Herbert I.

REGISTRATION NUMBER: 24,392

REFERENCE/DOCKET NUMBER: P-500-23744

TELEPHONE: (202) 887-0400

TELEFAX: (202) 835-0605

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 333 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-024-253-6

Query Match 55.0%; Score 33; DB 2; Length 333;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KFKKFKCK 10  
| | : | | |  
Db 199 KGKFKCK 206

## RESULT 15

US-08-714-070A-1

; Sequence 1, Application US/08714070A

; Patent No. 5834237

## GENERAL INFORMATION:

APPLICANT: JACOBS, Eric

APPLICANT: SILVESTRE, Nathalie

APPLICANT: SCHWEINBYRER, Ernst

TITLE OF INVENTION: COMBINED USE OF TWO EXPRESSION CASSETTES

TITLE OF INVENTION: FOR THE PRODUCTION OF A PROTEIN OF INTEREST

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

;; NUMBER OF SEQUENCES: 6  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: MORGAN & FINNEGAN  
;; STREET: 345 PARK AVENUE  
;; CITY: NEW YORK  
;; STATE: NEW YORK  
;; COUNTRY: USA  
;; ZIP: 10154  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: FLOPPY DISK, 3.5"  
;; COMPUTER: IBM PC COMPATIBLE  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: WORDPERFECT 5.1  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/647,714  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/388,809  
;; FILING DATE: 15-FEB-1995  
;; APPLICATION NUMBER: US 08/022,835  
;; FILING DATE: 25-FEB-1993  
;; APPLICATION NUMBER: US 07/599,491  
;; FILING DATE: 17-OCT-1990  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: LESLIE A. SERUNIAN  
;; REGISTRATION NUMBER: 35,353  
;; REFERENCE/DOCKET NUMBER: 2026-4092US2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 751-6840  
;; TELEFAX: (212) 751-6849  
;; TELEX: 421792  
;;  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 855 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
;; US-08-647-714-6

Query Match 56.7%; Score 34; DB 2; Length 855;  
Best Local Similarity 70.0%; Pred. No. 3e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KCKKFKFKCK 10  
||| ||| |  
Db 231 KCKDKKFKNGK 240

RESULT 11  
US-07-971-096-2  
; Sequence 2, Application US/07971096  
; Patent No. 5480972  
; GENERAL INFORMATION:  
; APPLICANT: Singh, Mohan Bir  
; APPLICANT: Avjoglu, Asil  
; APPLICANT: Knox, Robert Bruce  
; TITLE OF INVENTION: ALLERGENIC PROTEINS AND PEPTIDES FROM  
; TITLE OF INVENTION: JOHNSON GRASS POLLEN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0,  
; SOFTWARE: Version #1.25

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/971,096  
;; FILING DATE: 19921030  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mandragouras, Amy E.  
;; REGISTRATION NUMBER: 36,207  
;; REFERENCE/DOCKET NUMBER: IPC-042 (IMI-022)  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 261 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
;; US-07-971-096-2

Query Match 55.0%; Score 33; DB 1; Length 261;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KFKKFKCK 10  
|||: |||  
Db 156 KFRVVKCK 163

RESULT 12  
US-08-175-096-2  
; Sequence 2, Application US/08175096  
; Patent No. 5691167  
; GENERAL INFORMATION:  
; APPLICANT: Singh, Mohan Bir  
; APPLICANT: Avjoglu, Asil  
; APPLICANT: Knox, Robert Bruce  
; TITLE OF INVENTION: ALLERGENIC PROTEINS AND PEPTIDES FROM  
; TITLE OF INVENTION: JOHNSON GRASS POLLEN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0,  
; SOFTWARE: Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/175,096  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/971,096,  
; FILING DATE: OCT 30, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IPC-042 (IMI-022)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 261 amino acids



TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 855 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-022-835-6

Query Match 56.7%; Score 34; DB 1; Length 855;  
Best Local Similarity 70.0%; Pred. No. 3e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KCKFKKFKCK 10  
||| ||| |  
Db 231 KCKDKKFKNGK 240

RESULT 8  
US-08-388-809-6  
; Sequence 6, Application US/08388809  
; Patent No. 5576000  
; GENERAL INFORMATION:  
; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,  
; APPLICANT: GENOVEFFA; MARKHAM, PHILIP D. GALLO, ROBERT  
; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND  
; APPLICANT: GARTNER, SUZANNE  
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK, 3.5"  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/388,809  
FILING DATE: 15-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/022,835  
FILING DATE: 25-FEB-1993  
APPLICATION NUMBER: US 07/599,491  
FILING DATE: 17-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: LESLIE A. SERUNIAN  
REGISTRATION NUMBER: 35,353  
REFERENCE/DOCKET NUMBER: 2026-4092US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 855 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-388-809-6

Query Match 56.7%; Score 34; DB 1; Length 855;  
Best Local Similarity 70.0%; Pred. No. 3e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KCKFKKFKCK 10

Db 231 KCKDKKFKNGK 240  
||| ||| |

RESULT 9  
US-08-531-525-40  
; Sequence 40, Application US/08531525  
; Patent No. 5840683  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.  
; APPLICANT: Pincus, Matthew R.  
; APPLICANT: No. 5840683le, John F.  
; APPLICANT: Abajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
; TITLE OF INVENTION: Of P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: 'PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 184 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-531-525-40

Query Match 56.7%; Score 34; DB 2; Length 184;  
Best Local Similarity 66.7%; Pred. No. 88;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KCKFKKFKC 9  
||| ||| |  
Db 173 KCKKKKSC 181

RESULT 10  
US-08-647-714-6  
; Sequence 6, Application US/08647714  
; Patent No. 5869313  
; GENERAL INFORMATION:  
; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,  
; APPLICANT: GENOVEFFA; MARKHAM, PHILIP D. GALLO, ROBERT  
; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND  
; APPLICANT: GARTNER, SUZANNE  
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1  
; TITLE OF INVENTION: AND USES THEREOF

Db 3 KCKF-KFKC 10

RESULT 5

US-08-333-901-1

Sequence 1, Application US/08333901

Patent No. 559906

GENERAL INFORMATION:

APPLICANT: Dasmahapatra, Bimalendu

TITLE OF INVENTION: No. 559906el Protease Assays

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schering Plough Corporation

STREET: One Giralda Farms

CITY: Madison

STATE: New Jersey

COUNTRY: USA

ZIP: 07940

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 6.0.5

SOFTWARE: Microsoft Word 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/333,901

FILING DATE: 03-NOV-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/923,988

FILING DATE: 21-SEP-1992

APPLICATION NUMBER: PCT/US91/02283

FILING DATE: 10-APR-1991

APPLICATION NUMBER: 07/509,007

FILING DATE: 13-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Dulak, No. 559906man C.

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 881 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-333-901-1

Query Match 58.3%; Score 35; DB 1; Length 881;

Best Local Similarity 62.5%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9

Db 14 CRLKLLKC 21

RESULT 6

US-08-456-582-1

Sequence 1, Application US/08456582

Patent No. 5721133

GENERAL INFORMATION:

APPLICANT: Dasmahapatra, Bimalendu

TITLE OF INVENTION: No. 5721133el Protease Assays

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schering-Plough Corporation

STREET: One Giralda Farms

CITY: Madison

STATE: New Jersey

COUNTRY: USA

ZIP: 07940

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 6.0.5

SOFTWARE: Microsoft Word 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,582

FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/923,988

FILING DATE: 21-SEP-1992

APPLICATION NUMBER: PCT/US91/02283

FILING DATE: 10-APR-1991

APPLICATION NUMBER: 07/509,007

FILING DATE: 13-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Dulak, No. 5721133man C.

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 881 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-456-582-1

Query Match 58.3%; Score 35; DB 1; Length 881;

Best Local Similarity 62.5%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9

Db 14 CRLKLLKC 21

RESULT 7

US-08-022-835-6

Sequence 6, Application US/08022835

Patent No. 5420030

GENERAL INFORMATION:

APPLICANT: Reitz Jr., Marvin S.

APPLICANT: Franchini, Genoveffa

APPLICANT: Markham, Phillip D.

APPLICANT: Gallo, Robert C.

APPLICANT: Lori, Franco C.

APPLICANT: Popovic, Mikulas

APPLICANT: Garnter, Suzanne

TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES THEREOF

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DAREY & CUSHMAN

STREET: Eleventh Floor, 1615 L. Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036-5601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/022,835

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/599,491

FILING DATE: 17-OCT-1990

ATTORNEY/AGENT INFORMATION:

NAME: Scott, Watson T.

REGISTRATION NUMBER: 26,581

REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 861-3000

TELEFAX: (202) 822-0944

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 18:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; US-08-456-112B-18

Query Match 100.0%; Score 60; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCKFKFKCK 10  
|||||

Db 1 KCKFKFKCK 10

RESULT 3  
US-08-097-830E-15  
; Sequence 15, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.

;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-003  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; FEATURE:  
;; OTHER INFORMATION: sulfide bond between Cys and  
;; OTHER INFORMATION: Cys  
;; US-08-097-830E-15

Query Match 65.8%; Score 39.5; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 1.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KCKFKFKCK 9  
|||||

Db 3 KCKF-KFKC 10

RESULT 4  
US-08-456-112B-15  
; Sequence 15, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-15

Query Match 65.8%; Score 39.5; DB 2; Length 10;  
Best Local Similarity 88.9%; Pred. No. 1.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KCKFKFKCK 9  
|||||

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:38 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-18  
Perfect score: 60  
Sequence: 1 KCKFKKFKCK 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	10	1	US-08-097-830E-18
2	60	100.0	10	2	US-08-456-112B-18
3	39.5	65.8	10	1	US-08-097-830E-15
4	39.5	65.8	10	2	US-08-456-112B-15
5	35	58.3	881	1	US-08-333-901-1
6	35	58.3	881	1	US-08-456-582-1
7	34	56.7	855	1	US-08-022-835-6
8	34	56.7	855	1	US-08-388-809-6
9	34	56.7	184	2	US-08-531-525-40
10	34	56.7	855	2	US-08-647-714-6
11	33	55.0	261	1	US-07-971-096-2
12	33	55.0	261	1	US-08-175-096-2
13	33	55.0	333	2	US-08-436-463-6
14	33	55.0	333	2	US-08-024-253-6
15	32	53.3	775	2	US-08-714-070A-1
16	32	53.3	50	2	US-08-811-897A-7
17	32	53.3	476	2	US-08-811-897A-14
18	32	53.3	513	2	US-08-811-897A-15
19	32	53.3	448	2	US-08-811-897A-16
20	32	53.3	485	2	US-08-811-897A-17
21	32	53.3	448	2	US-08-811-897A-18
22	32	53.3	467	2	US-08-811-897A-19
23	32	53.3	476	2	US-08-811-897A-20
24	32	53.3	495	2	US-08-811-897A-21
25	32	53.3	448	2	US-08-811-897A-22
26	32	53.3	525	2	US-08-811-897A-23
27	32	53.3	476	2	US-08-811-897A-24
28	32	53.3	553	2	US-08-811-897A-25
29	32	53.3	475	2	US-08-811-897A-26
30	32	53.3	552	2	US-08-811-897A-27
31	32	53.3	476	2	US-08-811-897A-28
32	32	53.3	553	2	US-08-811-897A-29
33	32	53.3	376	2	US-08-465-976A-3
34	31.5	52.5	10	1	US-08-049-871-2
35	31.5	52.5	10	1	US-08-049-871-4
36	31.5	52.5	9	1	US-08-049-871-5
37	31.5	52.5	11	1	US-08-049-871-6
38	31.5	52.5	10	1	US-07-819-893-2
39	31.5	52.5	10	1	US-07-819-893-4

Sequence 5, Appli  
Sequence 6, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 5, Appli  
Sequence 50, Appli

ALIGNMENTS

RESULT 1  
US-08-097-830E-18  
; Sequence 18, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM: 3.50 inch, 1.44 Mb storage  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-18

Query Match 100.0%; Score 60; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KCKFKKFKCK 10  
| | | | | | | | | |  
Db 1 KCKFKKFKCK 10

RESULT 2  
US-08-456-112B-18  
; Sequence 18, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45

**THIS PAGE BLANK (USPTO)**

CC W21729-30). Compounds which interfere with the interaction of NuMA  
CC with a known NIP are used to modulate cell division and/or proliferation.  
CC AB, raised conventionally using NIP-1 or -2 as immunogen, are used to  
CC detect NIP (or their complexes) and to block their activity for  
CC diagnostic or therapeutic use, e.g. to detect defective NuMA or NIP  
CC which may be markers for aberrant (including malignant) cell growth  
CC (which can also be detected by nucleic acid sequencing). Also where  
CC malignancy is related to defects in NuMA or NIP, it can be treated by  
CC administration of the appropriate functional protein.  
SQ Sequence 2272 AA;

Query Match 58.3%; Score 35; DB 1; Length 2272;  
Best Local Similarity 62.5%; Pred. No. 6.8e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 CKFKKFKC 9  
I: ||||  
Db 14 CRLKKLKC 21

Search completed: September 7, 1999, 20:37:20  
Job time: 18469 sec

PT specific cell  
 PS Claim 7; Page 81-83; 106pp; English.  
 CC A multidomain protein (R95056) has a FLAG epitope, the DNA  
 CC binding domain of yeast GAL4, a portion of diphtheria toxin  
 CC that acts as a translocation domain, and a single chain antibody,  
 CC scFv, of monoclonal antibody FRP5 (raised against human tumour cell  
 CC HER2 antigen) that acts as a ligand domain. It is the product of a  
 CC fusion gene (729412), and is used with an effector nucleic acid that  
 CC comprises e.g. a gene to be delivered to a cell and a cognate  
 CC structure for the GAL4 DNA binding domain. This provides a novel  
 CC means of nucleic acid transfer, suitable for gene therapy.  
 SQ Sequence 615 AA;

Query Match 58.3%; Score 35; DB 1; Length 615;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9  
 I: | | | |  
 Db 28 CRLKLLKC 35

#### RESULT 13

W13082 ID W13082 standard; protein; 881 AA.

AC W13082; DT 08-MAY-1997 (first entry)  
 DE Yeast transcription activator GAL4.  
 KW Yeast; transcription; activator; GAL4; modification; binding; DNA;  
 KW hybrid regulatory protein; linker peptide; proteolytic cleavage;  
 KW identification; protease; inhibitor.  
 OS Saccharomyces cerevisiae.  
 FH Key Location/Qualifiers  
 FT peptide 1..147  
 FT /note= "DNA binding peptide"  
 FT peptide 788..881  
 FT /note= "DNA transcription modifying peptide"

US599906-A.

PD 04-FEB-1997.  
 PE 13-APR-1990; 509007.  
 PR 13-APR-1990; US-509007.  
 PR 10-APR-1991; WO-U02289.  
 PR 21-SEP-1992; US-923980.  
 PR 03-NOV-1994; US-333901.  
 PA (SCHE ) SCHERING CORP.  
 PI Dasmahapatra B;  
 DR WPI: 97-11837/11.  
 PT Hybrid regulatory protein - comprising DNA-binding peptide and DNA  
 PT transcription-modifying peptide, covalently joined by linker peptide  
 PT contg. protease cleavage site  
 PS Claim 5; Columns 15-20; 15pp; English.  
 CC The present sequence is the yeast transcription activator GAL4,  
 CC which can be used in the construction of a hybrid regulatory  
 CC protein (HRP), comprising the GAL4 DNA binding and transcription  
 CC modifying peptides covalently joined via a linker peptide. The HRP  
 CC is capable of altering the expression of gene(s) by binding to a  
 CC specific region of DNA controlling its expression, and is rendered  
 CC substantially inactive through proteolytic cleavage of the linker  
 CC peptide, i.e. the cleaved HRP has no more than 20 % of the  
 CC transcription activating activity of the intact HRP. The HRP can  
 CC be used to identify specific protease inhibitors.  
 SQ Sequence 881 AA;

Query Match 58.3%; Score 35; DB 1; Length 881;  
 Best Local Similarity 62.5%; Pred. No. 3e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9  
 I: | | | |  
 Db 14 CRLKLLKC 21

#### RESULT 14

W12096 ID W12096 standard; protein; 562 AA.

AC W12096; DT 30-SEP-1997 (first entry)  
 DE Adenovirus E1A associated protein BS69, inhibits replication.  
 DE Adenovirus; early region 1A; E1A; transactivation inhibitor; 243R;  
 KW 289R; viral replication inhibitor; immunoassay; gene therapy.  
 OS Adenovirus.  
 PN W09700323-A1.  
 PD 03-JAN-1997.  
 PF 14-JUN-1996; G01413.  
 PR 14-JUN-1995; GB-012092.  
 PA (VEHE-) VER HET NEDERLANDS KANKER INST.  
 PI Bernards R, Hateboer G;  
 DR WPI: 97-077529/07.  
 PT New adenovirus E1A associated protein BS69 - useful for inhibiting  
 PT trans:activation, e.g. as antiviral agent and for gene therapy of  
 PT tumours  
 PS Claim 1; Page 33-36; 49pp; English.  
 CC W12096 shows the adenovirus E1A-associated protein BS69. BS69 can  
 CC bind to early region 1A (E1A) proteins 243R and 289R to inhibit the  
 CC transactivation activity of 289R. In doing so the peptide also inhibits  
 CC viral replication and is therefore useful as an anti-adenoviral agent.  
 CC Vectors encoding the BS69 protein are also used in vivo for gene  
 CC therapy (e.g. to treat uncontrolled proliferation of tumour cells).  
 SQ Sequence 562 AA;

Query Match 58.3%; Score 35; DB 1; Length 562;  
 Best Local Similarity 70.0%; Pred. No. 2.1e+02;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KCKFKKFKCK 10  
 I: | | | |  
 Db 487 KCKQVKRCK 496

#### RESULT 15

W21731 ID W21731 standard; protein; 2272 AA.

AC W21731; DT 01-OCT-1997 (first entry)  
 DE GAL4/HA/NuMA fusion protein.  
 KW NIP-1; NIP-2; NuMA; nuclear mitotic apparatus; NuMA interacting protein;  
 KW cell division; proliferation; antibody; Ab; detection;  
 KW malignant cell growth.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT domain 1..147  
 FT /label= GAL4\_DNA\_binding\_domain  
 FT peptide 148..174  
 FT /label= Hemagglutinin\_epitope  
 FT protein 175..2272  
 FT /label= Residues\_18-2116\_of\_NuMA  
 FT region 365..1864  
 FT /label= Colled\_coil\_region  
 PN W09640917-A1.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; U09504.  
 PR 07-JUN-1995; US-478408.  
 PA (UYVA ) UNIV YALE.  
 PI McPherson SMG, Snyder MP;  
 DR WPI: 97-077270/07.  
 DR N-PSDB; T77782.  
 PT New nucleic acid encoding nuclear mitotic appts. Interacting  
 PT proteins - useful for modulating cell division and proliferation and  
 PT in diagnosis  
 PS Claim 14; Page 28-36; 78pp; English.  
 CC The sequences given in W21731-32 represent fusion proteins which contain  
 CC NuMA (nuclear mitotic apparatus). The fusion proteins were used in  
 CC the identification of NuMA interacting proteins (NIP's) (see also

PN WO9613599-A1.  
PD 09-MAY-1996.  
PF 31-OCT-1995; E04270.  
PR 01-NOV-1994; EP-810627.  
PA (WELS/) WELS W.  
PI Fominaya J, Wells W;  
DR WPI: 96-239505/24.  
DR N-PSDB; T29409.  
PT Nucleic acid transfer system for gene therapy, e.g. against cancer  
PT - includes toxin translocation domain to target nucleic acid to  
PT specific cell  
PS Claim 7; Page 64-65; 106pp; English.  
CC A multidomain protein (R95054) has a FLAG epitope, a portion  
CC of human transforming growth factor-alpha (TGF-a) that acts as a  
CC ligand domain, a non-cytotoxic portion of Pseudomonas aeruginosa  
CC exotoxin A acting as a translocation domain and the DNA  
CC binding domain of yeast GAL4. It is the product of a fusion  
CC gene (T29410) and can be expressed in E. coli (resulting in  
CC removal of an ompA signal peptide). It is used with an effector  
CC nucleic acid that comprises e.g. a gene to be delivered to  
CC a cell and a cognate structure for the GAL4 DNA binding domain.  
CC This provides a novel means of nucleic acid transfer, suitable  
CC for gene therapy.  
SQ Sequence 342 AA;

Query Match 58.3%; Score 35; DB 1; Length 342;  
Best Local Similarity 62.5%; Pred. No. 1.4e-02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CRFKFKFC 9  
I: || ||  
Db 201 CRUKLKC 208

RESULT 11  
R95055  
ID R95055 standard; Protein; 421 AA.  
AC R95055;  
DE 19-AUG-1996 (first entry)  
DE IL-2-DETA-DGAL4 multidomain protein.  
KW Nucleic acid transfer system; gene transfer; gene therapy;  
KW cell targeting; multidomain protein; vector; cancer;  
KW exotoxin A; DETA; ompA; signal peptide; GAL4; interleukin-2;  
KW IL-2.  
OS Chimeric synthetic;  
OS Chimeric Homo sapiens;  
OS Chimeric Pseudomonas aeruginosa;  
OS Chimeric Saccharomyces cerevisiae.  
FH Key Location/Qualifiers  
FT peptide 1..8  
FT /label= FLAG\_epitope  
FT peptide 9..17  
FT /label= Spacer  
FT domain 18..150  
FT /label= IL-2  
FT /note= "amino acids 1-113 of human IL-2  
FT 151  
FT /label= Spacer  
FT domain 152..266  
FT /label= ETA  
FT /note= "amino acids 252-366 of ETA"  
FT 267  
FT /label= Spacer  
FT domain 268..413  
FT /label= GAL4  
FT /note= "amino acids 2-147 of yeast GAL4"  
FT 414..421  
FT /label= Spacer  
FT /note= "endoplasmic reticulum retention signal"  
FT WO9613599-A1.  
PN 09-MAY-1996.  
PF 31-OCT-1995; E04270.

PR 01-NOV-1994; EP-810627.  
PA (WELS/) WELS W.  
PI Fominaya J, Wells W;  
DR WPI: 96-239505/24.  
DR N-PSDB; T29411.  
PT Nucleic acid transfer system for gene therapy, e.g. against cancer  
PT - includes toxin translocation domain to target nucleic acid to  
PT specific cell  
PS Claim 7; Page 67-69; 106pp; English.  
CC A multidomain protein (R95055) has a FLAG epitope, a portion  
CC of human interleukin-2 that acts as a ligand domain, a  
CC non-cytotoxic portion of Pseudomonas aeruginosa exotoxin A acting  
CC as a translocation domain and the DNA binding domain of yeast GAL4.  
CC It is the product of a fusion gene (T29411) and can be expressed  
CC in E. coli (resulting in removal of an ompA signal peptide). It is  
CC used with an effector nucleic acid that comprises e.g. a gene to be  
CC delivered to a cell and a cognate structure for the GAL4 DNA binding  
CC domain. This provides a novel means of nucleic acid transfer,  
CC suitable for gene therapy.  
SQ Sequence 421 AA;

Query Match 58.3%; Score 35; DB 1; Length 421;  
Best Local Similarity 62.5%; Pred. No. 1.6e-02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CRFKFKFC 9  
I: || ||  
Db 280 CRLLKLC 287

RESULT 12  
R95056  
ID R95056 standard; Protein; 615 AA.  
AC R95056;  
DE 19-AUG-1996 (first entry)  
DE GAL4-DT-scfv(FRP5) multidomain protein.  
KW Nucleic acid transfer system; gene transfer; gene therapy;  
KW cell targeting; multidomain protein; vector; cancer;  
KW GAL4; diphtheria toxin; single chain antibody; scFv; FRP5.  
OS Chimeric synthetic;  
OS Chimeric Saccharomyces cerevisiae;  
OS Chimeric Clostridium diphtheriae;  
OS Chimeric Mus sp.  
FH Key Location/Qualifiers  
FT peptide 1..8  
FT /label= FLAG\_epitope  
FT peptide 9..15  
FT /label= Spacer  
FT domain 16..161  
FT /label= GAL4  
FT /note= "amino acids 2-147 of yeast GAL4"  
FT 162..165  
FT /label= Spacer  
FT domain 166..354  
FT /label= DT  
FT /note= "amino acids 196-384 of diphtheria toxin"  
FT 355..361  
FT /label= Spacer  
FT domain 362..601  
FT /label= scfv(FRP5)  
FT peptide 602..615  
FT /label= Spacer  
FT WO9613599-A1.  
PN 09-MAY-1996.  
PF 31-OCT-1995; E04270.  
PR 01-NOV-1994; EP-810627.  
PA (WELS/) WELS W.  
PI Fominaya J, Wells W;  
DR WPI: 96-239505/24.  
DR N-PSDB; T29412.  
PT Nucleic acid transfer system for gene therapy, e.g. against cancer  
PT - includes toxin translocation domain to target nucleic acid to



FT peptide /note= "amino acids 196-384 of diphtheria toxin"  
 FT 355..363  
 FT /label= Spacer  
 FT domain 364..496  
 FT /label= IL-2  
 FT /note= amino acids 1-113 of human IL-2  
 PN WO9613599-A1.  
 PD 09-MAY-1996.  
 PR 31-OCT-1995; E04270.  
 PR 01-NOV-1994; EP-810627.  
 PA (WELS/) WELS W.  
 PI Fominaya J. Wels W;  
 DR WPI: 96-239505/24.  
 DR N-PSDB; T29414.  
 CC A multidentin protein (R95058) has a FLAG epitope, the DNA  
 CC binding domain of yeast GAL4, a portion of diphtheria toxin  
 CC that acts as a translocation domain, and an interleukin-2 moiety  
 CC that acts as a ligand domain. It is the product of a fusion gene  
 CC (T29414), and is used with an effector nucleic acid that  
 CC comprises e.g. a gene to be delivered to a cell and a cognate  
 CC structure for the GAL4 DNA binding domain. This provides a novel  
 CC means of nucleic acid transfer, suitable for gene therapy.  
 SQ Sequence 496 AA;

Query Match 58.3%; Score 35; DB 1; Length 496;  
 Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9  
 I: || ||  
 Db 28 CRUKKLKC 35

RESULT 9  
 ID R95053 standard; Protein: 530 AA.  
 AC R95053;  
 DT 18-AUG-1996 (first entry)  
 DE scfv(FRP5)-DETA-DGAL4 multidomain protein.  
 KW Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer;  
 KW exotoxin A; DETA; single chain antibody; scFv; Gal4.  
 OS Chimeric synthetic;  
 OS Chimeric Mus sp.;  
 OS Chimeric Pseudomonas aeruginosa;  
 OS Chimeric Saccharomyces cerevisiae.  
 FH Key Location/Qualifiers  
 FT peptide 1..8  
 FT /label= FLAG\_epitope  
 FT peptide 9..117  
 FT /label= Spacer  
 FT domain 18..257  
 FT /label= scFv(FRP5)  
 FT peptide 258..260  
 FT /label= Spacer  
 FT domain 261..375  
 FT /label= ETA  
 FT /note= "amino acids 252-366 of exotoxin-A"  
 FT peptide 376  
 FT /label= Spacer  
 FT domain 377..522  
 FT /label= GAL4  
 FT /note= "amino acids 2-147 of yeast GAL4"  
 FT peptide 523..530  
 FT /label= Spacer  
 FT /note= "endoplasmic reticulum retention peptide"  
 PN WO9613599-A1.  
 PD 09-MAY-1996.

PF 31-OCT-1995; E04270.  
 PR 01-NOV-1994; EP-810627.  
 PA (WELS/) WELS W.  
 PI Fominaya J. Wels W;  
 DR WPI: 96-239505/24.  
 DR N-PSDB; T29409.  
 CC Nucleic acid transfer system for gene therapy, e.g. against cancer  
 CC - includes toxin translocation domain to target nucleic acid to  
 CC specific cell  
 PS Claim 7; Page 59-61; 106pp; English.  
 CC A multidomain protein (R95053) has a FLAG epitope, a single  
 CC chain antibody, scFv, of monoclonal antibody FRP5 (raised  
 CC against human tumour cell HER2 antigen) that acts as a ligand  
 CC domain, a non-cytotoxic portion of Pseudomonas aeruginosa  
 CC exotoxin A acting as a translocation domain and the DNA  
 CC binding domain of yeast GAL4. It is the product of a fusion  
 CC gene (T29409) and can be expressed in E. coli (resulting in  
 CC removal of ompA signal peptide). It is used with an effector  
 CC nucleic acid that comprises e.g. a gene to be delivered to  
 CC a cell and a cognate structure for the GAL4 DNA binding domain.  
 CC This provides a novel means of nucleic acid transfer, suitable  
 CC for gene therapy.  
 SQ Sequence 530 AA;

Query Match 58.3%; Score 35; DB 1; Length 530;  
 Best Local Similarity 62.5%; Pred. No. 2e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9  
 I: || ||  
 Db 389 CRUKKLKC 396

RESULT 10  
 ID R95054 standard; Protein: 342 AA.  
 AC R95054;  
 DT 19-AUG-1996 (first entry)  
 DE TGF-a-DETA-DGAL4 multidomain protein.  
 KW Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer;  
 KW exotoxin A; DETA; ompA; signal peptide; GAL4; TGF-a;  
 KW transforming growth factor-alpha.  
 OS Chimeric Homo sapiens;  
 OS Chimeric Pseudomonas aeruginosa;  
 OS Chimeric Saccharomyces cerevisiae.  
 FH Key Location/Qualifiers  
 FT peptide 1..8  
 FT /label= FLAG\_epitope  
 FT peptide 9..12  
 FT /label= Spacer  
 FT domain 13..62  
 FT /label= TGF-a  
 FT /note= "amino acids 1-50 of human TGF-a"  
 FT peptide 63..65  
 FT /label= Spacer  
 FT peptide 66..71  
 FT /label= Hexa-histidine  
 FT peptide 72  
 FT /label= Spacer  
 FT domain 73..187  
 FT /label= ETA  
 FT /note= "amino acids 252-366 of ETA"  
 FT peptide 188  
 FT /label= Spacer  
 FT domain 189..334  
 FT /label= GAL4  
 FT /note= "amino acids 2-147 of yeast GAL4"  
 FT peptide 335..342  
 FT /label= Spacer  
 FT /note= "endoplasmic reticulum retention signal"  
 FT

Query Match 58.3%; Score 35; DB 1; Length 882;  
 Best Local Similarity 62.5%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9  
 I: ||||  
 Db 14 CRLKLLKC 21

RESULT 6

R80490 ID R80490 standard; Protein; 169 AA.  
 AC R80490;  
 DT 26-JAN-1996 (first entry)  
 DE GAL4 DNA binding domain.  
 KW DNA binding domain; GAL4-receptor; steroid/thyroid receptor protein;  
 OS Yeast.  
 PN WO9518380-A.  
 PD 08-JUL-1995.  
 PF 29-DEC-1993; U14426.  
 PR 30-DEC-1993; US-177740.  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PI Blumberg B, Evans RM, Rangarajan PN, Umesono K;  
 DR WPI: 95-246506/32.  
 DT N-PSDB: Q99220.  
 PT GAL4 receptor constructs formed from steroid/thyroid receptor proteins  
 (STRP) - useful in assays to determine the ligand for an orphan member of  
 the STRP super-family  
 PS Claim 2; Fig 1; 49pp; English.  
 CC A method is claimed to determine the ligand or ligand precursor for  
 an orphan member of the steroid/thyroid receptor protein superfamily.  
 CC The method comprises contacting cells contg. a modified form of the  
 CC orphan member with a putative ligand. The modified form of the  
 CC orphan member is produced by introducing the DNA binding domain of  
 CC GAL4 into it. The cells contain a GAL4 response element operatively  
 CC linked to a receptor gene. The process is then monitored by  
 CC monitoring the expression of the reporter gene product. The DNA  
 CC binding domain of GAL4 pref. comprises R80490 residues 1-147. Q99220  
 CC is the nt sequence of avector useful for the prepn. of GAL4-receptor  
 CC fusion proteins, including an exemplary segment encoding a GAL4 DNA  
 CC binding domain useful for the method of the invention.  
 CC Sequence 169 AA;  
 SQ

Query Match 58.3%; Score 35; DB 1; Length 169;  
 Best Local Similarity 62.5%; Pred. No. 75;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9  
 I: ||||  
 Db 14 CRLKLLKC 21

RESULT 7

R95057 ID R95057 standard; Protein; 617 AA.  
 AC R95057;  
 DT 19-AUG-1996 (first entry)  
 DE GAL4-DT-scfv(FRP5) multidomain protein.  
 KW Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer;  
 KW GAL4; diphtheria toxin; single chain antibody; scfv; FRP5.  
 OS Chimeric synthetic;  
 OS Chimeric Saccharomyces cerevisiae;  
 OS Chimeric Clostridium diphtheriae;  
 OS Chimeric Mys sp. Location/Qualifiers  
 FH Key 1  
 FT Peptide 1  
 FT Peptide 2  
 FT Peptide 2  
 FT Peptide 10  
 FT Peptide 10

FT domain /label= Spacer  
 FT 18..163  
 FT /label= GAL4  
 FT /note= "amino acids 2-147 of yeast GAL4"  
 FT 164..167  
 FT /label= Spacer  
 FT 168..356  
 FT /label= DT  
 FT /note= "amino acids 196-384 of diphtheria toxin"  
 FT 357..363  
 FT /label= Spacer  
 FT 364..603  
 FT /label= scfv(FRP5)  
 FT 604..617  
 FT /label= Spacer  
 PN WO9613599-A1.  
 PD 09-MAY-1996.  
 PF 31-OCT-1995; E04270.  
 PR 01-NOV-1994; EP-810627.  
 PA (WELS/) WELS W.  
 PI Fominaya J, Wells W;  
 DR WPI: 96-239505/24.  
 DT N-PSDB: T29413.  
 PT Nucleic acid transfer system for gene therapy, e.g. against cancer  
 PT - includes toxin translocation domain to target nucleic acid to  
 PT specific cell  
 PS Claim 7; Page 87-89; 106pp; English.  
 CC A multidomain protein (R95057) comprises a FLAG epitope, the DNA  
 CC binding domain of yeast GAL4, a portion of diphtheria toxin  
 CC that acts as a translocation domain, and a single chain antibody,  
 CC scfv, of monoclonal antibody FRP5 (raised against human tumour cell  
 CC HER2 antigen) that acts as a ligand domain. It is the product of a  
 CC fusion gene (T29413), and is used with an effector nucleic acid that  
 CC comprises e.g. a gene to be delivered to a cell and a cognate  
 CC structure for the GAL4 DNA binding domain. This provides a novel  
 CC means of nucleic acid transfer, suitable for gene therapy.  
 SQ Sequence 617 AA;

Query Match 58.3%; Score 35; DB 1; Length 617;  
 Best Local Similarity 62.5%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKKFKC 9  
 I: ||||  
 Db 30 CRLKLLKC 37

RESULT 8

R95058 ID R95058 standard; Protein; 496 AA.  
 AC R95058;  
 DT 19-AUG-1996 (first entry)  
 DE GAL4-DT-IL-2 multidomain protein.  
 KW Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer;  
 KW GAL4; diphtheria toxin; interleukin-2; IL-2.  
 OS Chimeric synthetic;  
 OS Chimeric Saccharomyces cerevisiae;  
 OS Chimeric Clostridium diphtheriae;  
 OS Chimeric Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..8  
 FT Peptide 9  
 FT Peptide 9  
 FT Peptide 16  
 FT domain /label= Spacer  
 FT 16..161  
 FT /label= GAL4  
 FT /note= "amino acids 2-147 of yeast GAL4"  
 FT 162..165  
 FT /label= Spacer  
 FT 166..354  
 FT /label= DT

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 60; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KCKFKFKFC 10  
 |||||  
 Db 1 KCKFKFKFC 10  
 |||||

RESULT 3

R71786  
 ID R71786 standard; peptide; 10 AA.  
 AC R71786;  
 DE 01-OCT-1995 (first entry)  
 DT Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M.  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 17: Page 21: 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 65.8%; Score 39.5; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KCKFKFKFC 9  
 |||||  
 Db 3 KCKF-KFKC 10

RESULT 4

W21603  
 ID W21603 standard; peptide; 10 AA.  
 AC W21603;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #15.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M; Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 19; Page 25; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 65.8%; Score 39.5; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KCKFKFKFC 9  
 |||||  
 Db 3 KCKF-KFKC 10

RESULT 5

R51930  
 ID R51930 standard; Protein; 882 AA.  
 AC R51930;  
 DT 24-NOV-1994 (first entry)  
 DE Mutated GAL4 protein with Asp inserted between Asp863 and Val864.  
 KW Mutated GAL4; heterologous gene expression; enhanced; increased;  
 KW protein production; positive regulator; transcription activator;  
 KW galactose metabolism.  
 OS Eukaryota.  
 FH Key Location/Qualifiers  
 FT misc\_difference 64  
 FT misc\_difference 233  
 FT misc\_difference 459  
 FT misc\_difference 664  
 FT region  
 FT J06078767-A.  
 PN 22-MAR-1994.  
 PF 07-SEP-1992; 262723.  
 PR 07-SEP-1992; JP-262723.  
 PA (GREEN) GREEN CROSS CORP.  
 DR WPI; 94-131274/16.  
 DR N-PSDB; Q61607.  
 PT Mutated GAL4 and a method for the expression of foreign protein -  
 PT useful for enhancing expression of protein  
 PS Claim 1; Page 10-12; 20pp; Japanese.  
 CC GAL4 in which an Asp residue is inserted between wldt-type Asp863  
 CC and Val864 is claimed. The mutated GAL4 can be used for increasing  
 CC the expression of foreign genes under the control of the GAL1, GAL7  
 CC or GAL10 promoters.  
 SQ Sequence 882 AA;

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:19 ; Search time 147.16 seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-18  
Perfect score: 60  
Sequence: 1 KCKFKKFKCK 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	10	1 R71789	Peptide neutralisi
2	60	100.0	10	1 W21606	Antibiotic potenti
3	39.5	65.8	10	1 R71786	Peptide neutralisi
4	39.5	65.8	10	1 W21603	Antibiotic potenti
5	35	58.3	882	1 R51930	Mutated GAL4 prote
6	35	58.3	169	1 R80490	GAL4 DNA binding d
7	35	58.3	617	1 R95057	GAL4-DT-scrv(FRP5)
8	35	58.3	496	1 R95058	GAL4-DT-IL-2 multi
9	35	58.3	530	1 R95059	SCFV(FRP5)-DETA-DG
10	35	58.3	342	1 R95054	TGF-a-DETA-DGALA m
11	35	58.3	421	1 R95055	IL-2-DETA-DGALA mu
12	35	58.3	615	1 R95056	GAL4-DT-scrv(FRP5)
13	35	58.3	881	1 W13082	Yeast transcriptio
14	35	58.3	562	1 W12096	Adenovirus E1A ass
15	35	58.3	2772	1 W21731	GAL4/HA/NUMA fusio
16	35	58.3	226	1 W14300	GAL4 DNA binding d
17	35	58.3	881	1 W47124	S. cerevisiae tran
18	35	58.3	378	1 W64589	Murine cytomagalov
19	34	56.7	855	1 R14905	HIV-1 BA-L clone-e
20	34	56.7	855	1 W11581	Human immunodefici
21	34	56.7	185	1 W03547	ORF-4 protein sequ
22	34	56.7	557	1 W25258	Human TRAF5, a nov
23	34	56.7	557	1 W27610	Human TRAF5, TRAF5
24	34	56.7	498	1 W37055	HIV-1 breakthrough
25	34	56.7	486	1 W37056	HIV-1 breakthrough
26	34	56.7	491	1 W37057	HIV-1 breakthrough
27	34	56.7	859	1 W43080	Env protein of the
28	34	56.7	855	1 W88113	Env protein of the
29	33	55.0	1456	1 R24033	Soluble mannose re
30	33	55.0	333	1 R11014	Feline immunoglobu
31	33	55.0	333	1 R48367	Constant region of
32	33	55.0	13	1 R54971	Sorhi grass pollen
33	33	55.0	261	1 R54949	Johnson grass poll
34	33	55.0	333	1 R54094	Sequence encoded b
35	32	53.3	3080	1 R35081	ZYMV polyprotein.
36	32	53.3	476	1 R58669	Human PACAP recept
37	32	53.3	485	1 R58657	Bovine PACAP recep
38	32	53.3	448	1 R58665	Bovine PACAP recep
39	32	53.3	475	1 R58670	Human PACAP recept
40	32	53.3	448	1 R58666	Rat PACAP receptor
41	32	53.3	467	1 R58671	Human PACAP recept
42	32	53.3	467	1 R58656	Rat PACAP receptor
43	32	53.3	495	1 R58658	Rat PACAP receptor

44 32 53.3 513 1 R58655 Bovine PACAP recep  
45 32 53.3 476 1 R58667 Rat PACAP receptor

## ALIGNMENTS

RESULT 1  
R71789  
ID R71789 standard; peptide; 10 AA.  
AC R71789;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 2..9  
PN W09503327-A.  
PF 02-FEB-1995.  
PR 21-JUL-1994; E02413.  
PA 26-JUL-1993; US-097830.  
PI (BIOS-) BIOSYNTH SRL.  
DR Porro M;  
WPI: 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids  
PS Claim 20; Page 21; 26pp; English.  
CC New Peptides are claimed which are linear or cyclic peptides which include units of formula: (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new peptides. 10 AA;  
SQ Sequence 10 AA;

Query Match 100.0%; Score 60; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KCKFKKFKCK 10  
| | | | | | | | | |  
Db 1 KCKFKKFKCK 10

RESULT 2  
W21606  
ID W21606 standard; peptide; 10 AA.  
AC W21606;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #18.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 2..9  
PN W09638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
DR Porro M, Varria M;  
WPI: 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic required  
PS Claim 22; Page 26; 37pp; English.

**THIS PAGE BLANK (USPTO)**

## RESULT 13

O34181 PRELIMINARY; PRT; 99 AA.  
 ID O34181;  
 AC O34181;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
 DE OLIGOPETIDE ABC TRANSPORTER (FRAGMENT).  
 OS Halobacterium volcanii (Haloflex volcanii).  
 OG Plasmid pHV3.  
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloflex.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DS2;  
 RA FARAHANI R., IMBEAULT J.-C., ST JEAN A., CHAN C.C.-Y., ALLARD G.,  
 RA CHARLEBOIS R.L.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U95372; AAB71803.1; -;  
 KW Plasmid.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 99 AA; 10635 MW; C1D985A8 CRC32;

Query Match 59.3%; Score 35; DB 1; Length 99;

Best Local Similarity 66.7%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TRCRFKRRC 10  
 | | | | |  
 DB 33 TGCRFHTRC 41

## RESULT 14

O86132 PRELIMINARY; PRT; 335 AA.  
 ID O86132;  
 AC O86132;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE ORNITHINE CARBAMOYLTRANSFERASE (EC 2.1.3.3).  
 GN ARCB.  
 OS Bacillus licheniformis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 14580;  
 RA MAGHNOUJ A., FRANCO DE SOUSA CABRAL T., STALON V., VANDER WAUVEN C.;  
 RT "The arginine deiminase pathway in Bacillus licheniformis: the genes  
 RT and their regulation by the ArgR protein.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y17554; CAA76778.1; -;  
 DR PFAM; PF00185; OTCace; 1.  
 KW Transferase.  
 SQ SEQUENCE 335 AA; 37659 MW; B2205EE7 CRC32;

Query Match 59.3%; Score 35; DB 2; Length 335;

Best Local Similarity 60.0%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RTRCRFKRRC 10  
 | | | | |  
 DB 61 RTRCAFTTAC 70

## RESULT 15

O15504 PRELIMINARY; PRT; 423 AA.  
 ID O15504;  
 AC O15504;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)

DE CG1.  
 GN CG1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA VAN LAER L., VAN CAMP G., VAN ZUYLEN D., GREEN E., VERSTREKEN M.,  
 RA SCHATTEMAN I., VAN DE HEYNING P., BALEMANS W., COUCKE P.,  
 RA GREINWALD J.H., SMITH R.J.H., HUIZING E., WILLEMS P.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U97198; AAC33794.1; -;  
 SQ SEQUENCE 423 AA; 44871 MW; 373E13CC CRC32;

Query Match 59.3%; Score 35; DB 4; Length 423;

Best Local Similarity 75.0%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRFKRRC 10  
 | | | | |  
 DB 11 RCRFGDRC 18

Search completed: September 7, 1999, 20:34:50  
 Job time: 19746 sec

```

RESULT 9
O93058 ID O93058 PRELIMINARY; PRT; 1616 AA.
AC O93058;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE 180K PROTEIN.
OS Tobacco mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OM;
RX MEDLINE; 82239259.
RA MESHI T., OHNO T., OKADA Y.;
RT "Nucleotide sequence and its character of cistron coding for the 30 K
protein of tobacco mosaic virus (OM strain).";
RL J. Biochem. 91:1441-1444(1982).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OM;
RX MEDLINE; 84029158.
RA MESHI T., ISHIKAWA M., TAKAMATSU N., OHNO T., OKADA Y.;
RT "The 5'-terminal sequence of TMV RNA. Question on the polymorphism
found in vulgar strain.";
RL FEBS Lett. 162:282-285(1983).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=OM;
RA WATANABE T., HIBI T., ISHIHAMA A.;
RT "Nucleotide sequence of the coding region for 180K protein of tobacco
mosaic virus common strain OM.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR ENBL; D78608; BA11429.1; -.
DR PFAM; PF00978; RNA_dep_RNAPol2; 1.
DR PFAM; PF01443; Viral_helicase1; 1.
SQ SEQUENCE 1616 AA; 183330 MW; 9638E8EC CRC32;

Query Match 61.0%; Score 36; DB 12; Length 1616;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TRCRFKR 8
DB 896 TRCQFKR 902
|||||

RESULT 10
O41340 ID O41340 PRELIMINARY; PRT; 1117 AA.
AC O41340;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE 130K PROTEIN.
OS Tobacco mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RAKKYO;
RX MEDLINE; 96265021.
RA CHEN J., WATANABE Y., SAKO N., OHSIMA K., OKADA Y.;
RT "Complete nucleotide sequence and synthesis of infectious in vitro
transcripts from a full-length cDNA clone of a rakkyo strain of
tobacco mosaic virus.";
RL Arch. Virol. 141:885-900(1996).
DR ENBL; D63809; BAA09877.1; -.
DR PFAM; PF01443; Viral_helicase1; 1.
SQ SEQUENCE 1117 AA; 126217 MW; 9A9FCB0A CRC32;

Query Match 61.0%; Score 36; DB 12; Length 1616;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TRCRFKR 8
DB 896 TRCQFKR 902
|||||

```

```

Query Match 61.0%; Score 36; DB 12; Length 1117;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TRCRFKR 8
DB 896 TRCQFKR 902
|||||

RESULT 11
Q98745 ID Q98745 PRELIMINARY; PRT; 1616 AA.
AC Q98745;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE 180K PROTEIN.
OS Tobacco mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RAKKYO;
RX MEDLINE; 96265021.
RA CHEN J., WATANABE Y., SAKO N., OHSIMA K., OKADA Y.;
RT "Complete nucleotide sequence and synthesis of infectious in vitro
transcripts from a full-length cDNA clone of a rakkyo strain of
tobacco mosaic virus.";
RL Arch. Virol. 141:885-900(1996).
DR ENBL; D63809; BAA09876.1; -.
DR PFAM; PF00978; RNA_dep_RNAPol2; 1.
DR PFAM; PF01443; Viral_helicase1; 1.
SQ SEQUENCE 1616 AA; 183579 MW; 54660605 CRC32;

Query Match 61.0%; Score 36; DB 12; Length 1616;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TRCRFKR 8
DB 896 TRCQFKR 902
|||||

RESULT 12
Q9YQ08 ID Q9YQ08 PRELIMINARY; PRT; 1616 AA.
AC Q9YQ08;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE 183 KDA PROTEIN.
OS Tobacco mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B935A;
RA XUE C., ZHOU X., CHEN Q., QI Y., LI D.;
RT "Complete nucleotide sequence and genome organization of tobacco
mosaic virus isolated from Vicia faba.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR ENBL; AJ011933; CAA09875.1; -.
SQ SEQUENCE 1616 AA; 183297 MW; 65551E8F CRC32;

Query Match 61.0%; Score 36; DB 12; Length 1616;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TRCRFKR 8
DB 896 TRCQFKR 902
|||||

```

DT	01-AUG-1998 (TremBrel. 07, Last sequence update)
DE	01-AUG-1998 (TremBrel. 07, Last annotation update)
DD	F27L4.13 PROTEIN.
DN	F27L4.13.
OS	Arabidopsis thaliana (Mouse-ear cross).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC	core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC	Arabidopsids.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. COLUMBIA;
RA	ROUNSELEY S.D., KAUL S., LIN X., KETCHUM K.A., CROSBY M.L.,
RA	BRANDON R.C., SYKES S.M., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,
RA	SOMERVILLE C.R., VENTER J.C.;
RL	Submitted (May-1998) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AC004482; AAC17093.1; -
SQ	SEQUENCE 278 AA; 30944 MW; FF314E78 CRC32;
Query Match	61.0%; Score 36; DB 10; Length 278;
Best Local Similarity	50.0%; Pred.No.16;
Matches	5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY	1 RTCRCKRRKC 10
	:: ::
DB	109 RDCKRCKC 118
RESULT	8
Q88561	
ID	PRELIMINARY; PRT; 1117 AA.
AC	Q88561;
DT	01-NOV-1996 (TremBrel. 01, Created)
DD	01-NOV-1996 (TremBrel. 01, Last sequence update)
DT	01-MAY-1999 (TremBrel. 10, Last annotation update)
DE	130K PROTEIN.
OS	Tobacco mosaic virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=OM;
RA	MEDLINE: 82239259.
RX	MESH T., OHNO T., OKADA Y.;
RT	"Nucleotide sequence and its character of cistron coding for the 3
RT	protein of tobacco mosaic virus (OM strain).";
RL	J. Biochem. 91:1441-1444(1982).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=OM;
RX	MEDLINE: 84029158.
RX	MESH T., ISHIKAWA M., TAKAMATSU N., OHNO T., OKADA Y.;
RT	"The 5'-terminal sequence of TMV RNA. Question on the polymorphism
RT	found in vulgaris strain.";
RL	FEBS Lett. 162:282-285(1983).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=OM;
RA	WATANABE T., HIBI T., ISHIHAMA A.;
RL	Submitted (DPC-1995) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; D78608; BAAL1430.1; -
DR	PFAM; PF01443; Viral_helicase1.1.
SQ	SEQUENCE 1117 AA; 126168 MW; FFF2883F CRC32;
Query Match	61.0%; Score 36; DB 12; Length 1117;
Best Local Similarity	85.7%; Pred.No.46;
Matches	6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



```

Db 57 RYKCRFDRKC 66

RESULT 2
ID Q84495 PRELIMINARY; PRT; 69 AA.
AC Q84495;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE FRARCA 2X.
GN A175R.
OS Paramacium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95133167.
RA LU Z., LI Y., ZHANG Y., KUTISH G.F., ROCK D.L., VAN ERTEN J.L.;
RT "Analysis of 45 kb of DNA located at the left end of the chlorella
RT virus PBCV-1 genome.";
RL Virology 206:339-352(1995).
DR EMBL; U42580; AAC96543.1; -.
SQ SEQUENCE 59 AA; 8091 MW; B6C2535B CRC32;

Query Match 66.1%; Score 39; DB 12; Length 69;
Best Local Similarity 60.0%; Pred. No. 1.5;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RTRCRFRRC 10
| | | | |
DB 36 RARCAFRAC 45

RESULT 3
ID Q44456 PRELIMINARY; PRT; 606 AA.
AC Q44456;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE C09G12.6 PROTEIN.
GN C09G12.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX LE T.T., HARPER M.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF038608; AAC25817.1; -.
SQ SEQUENCE 606 AA; 71785 MW; 008BCD97 CRC32;

Query Match 64.4%; Score 38; DB 5; Length 606;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TRCRFRRC 10
| | | | |
DB 134 TSCRFRPRC 142

RESULT 4
ID O53395 PRELIMINARY; PRT; 1538 AA.
AC O53395;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE PGRS-FAMILY PROTEIN.
GN PELGRS.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE; 98295987.
RA COLE S.T., BROSCHE R., PARKHILL J., GARNIER T., CHURCHER C., HARRIS D.,
RA GORDON S.V., EIGLMEIER K., GAS S., BARRY III C.E., TEKAIA F.,
RA BADCOCK K., BASHAM D., BROWN D., CHILLINGWORTH T., CONNOR R.,
RA DAVIES R., DEVLIN K., FELTWELL T., GENTLES S., HAMLIN N., HOLROYD S.,
RA HORNSBY T., JAGELS K., KROGH A., MCLEAN J., MOULE S., MURPHY L.,
RA OLIVER S., OSBORNE J., QUAIL M.A., RAJANDREAM M.A., ROGERS J.,
RA RUTTER S., SEGER K., SKELTON S., SQUARES S., SQUARES R., SULSTON J.E.,
RA TAYLOR K., WHITEHEAD S., BARRELL B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX PARKHILL J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL021841; CAAL17117.1; -.
DR PFAM; PF00934; PE.1.
SQ SEQUENCE 1538 AA; 129387 MW; 019442C9 CRC32;

Query Match 61.0%; Score 36; DB 2; Length 1538;
Best Local Similarity 60.0%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RTRCRFRRC 10
| | | | |
DB 1473 RRRCWRQRC 1482

RESULT 5
ID O45666 PRELIMINARY; PRT; 477 AA.
AC O45666;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE NHR-49 PROTEIN.
GN NHR-49.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:48 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-17

Perfect score: 59

Sequence: 1 RTRCRFKRRC 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	40	67.8	Q19496	Q19496 caenorhabdi
2	39	66.1	Q84495	Q84495 paramyosin
3	38	64.4	Q44456	Q44456 caenorhabdi
4	36	61.0	Q53395	Q53395 mycobacteri
5	36	61.0	Q45666	Q45666 caenorhabdi
6	36	61.0	Q39682	Q39682 daucus caro
7	36	61.0	Q84833	Q84833 arabidopsis
8	36	61.0	Q88561	Q88561 tobacco mos
9	36	61.0	Q93058	Q93058 tobacco mos
10	36	61.0	Q41340	Q41340 tobacco mos
11	36	61.0	Q98745	Q98745 tobacco mos
12	36	61.0	Q98708	Q98708 tobacco mos
13	35	59.3	Q34181	Q34181 halobacteri
14	35	59.3	Q86132	Q86132 bacillus li
15	35	59.3	Q41504	Q41504 homo sapien
16	35	59.3	Q77099	Q77099 aedes aegypt
17	35	59.3	Q77100	Q77100 aedes aegypt
18	35	59.3	Q28226	Q28226 canis famli
19	35	59.3	Q54845	Q54845 mus musculu
20	35	59.3	Q88623	Q88623 mus musculu
21	34	57.6	Q59021	Q59021 methanococ
22	34	57.6	Q28504	Q28504 archaeoglob
23	34	57.6	Q86315	Q86315 mycobacteri
24	34	57.6	Q75604	Q75604 homo sapien
25	34	57.6	Q75988	Q75988 homo sapien
26	34	57.6	Q16783	Q16783 caenorhabdi
27	34	57.6	Q4365	Q4365 caenorhabdi
28	34	57.6	Q46175	Q46175 bombyx mori
29	34	57.6	Q99212	Q99212 echinococcu

30	34	57.6	235	5	Q22621	Q22621 caenorhabdi
31	34	57.6	649	5	Q22225	Q22225 caenorhabdi
32	34	57.6	46	11	Q64412	Q64412 cavia porce
33	34	57.6	486	12	P87724	P87724 group a rot
34	33	55.9	55	2	O54490	O54490 staphylococ
35	33	55.9	234	2	P72305	P72305 rhodococcus
36	33	55.9	330	2	Q82128	Q82128 treponema d
37	33	55.9	489	3	O60062	O60062 schizosacch
38	33	55.9	1224	4	P78311	P78311 homo sapien
39	33	55.9	1265	4	P78312	P78312 homo sapien
40	33	55.9	792	4	P78313	P78313 homo sapien
41	33	55.9	732	4	O43607	O43607 homo sapien
42	33	55.9	369	4	Q06246	Q06246 homo sapien
43	33	55.9	508	4	Q02930	Q02930 homo sapien
44	33	55.9	338	5	O01899	O01899 caenorhabdi
45	33	55.9	971	5	O45785	O45785 caenorhabdi

ALIGNMENTS

RESULT 1  
Q19496 PRELIMINARY: PRT: 548 AA.  
ID Q19496;  
AC Q19496;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
DE HYPOTHETICAL 63.5 KD PROTEIN.  
GN F16H11.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COOPER J., COULSON A.,  
RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., DUBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KESHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATSON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPRAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans";  
RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COOPER J., COULSON A.,  
RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., DUBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KESHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATSON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPRAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans";  
RL Nature 368:32-38(1994).  
RN [3]  
RP Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

Query Match 67.8%; Score 40; DB 5; Length 548;  
Best Local Similarity 60.0%; Pred. No. 5;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RTRCRFKRRC 10  
| :|:| | |

**THIS PAGE BLANK (USPTO)**

RL PROC. NATL. ACAD. SCI. U.S.A. 84:5316-5319(1987).

RP [4] SEQUENCE FROM N.A.  
 RX MEDLINE; 88181903.  
 RA HECHT N.B.;  
 RT "Gene expression during spermatogenesis.";  
 RL ANN. N.Y. ACAD. SCI. 513:90-101(1987).

RN [5] SEQUENCE FROM N.A.

RP STRAIN-C129;

RC SCHLUETER G., ENGEL W.;

RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBT DATA BANKS.

CC -!- FUNCTION: PROTAMINES SUBSTITUTES FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.

CC -!- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE

CC DNA-HELIX (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: NUCLEAR.

CC -!- TISSUE SPECIFICITY: TESTIS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; K02926; G200489; -

DR EMBL; X07625; G53789; -

DR EMBL; X14003; G53791; -

DR EMBL; M27500; G200503; -

DR EMBL; Z47352; E245955; -

DR PIR; A02660; HSMSS1.

DR PIR; S03820; S03820.

DR PIR; A28331; A28331.

DR MGI; MGI:37765; PRM1.

DR PROSITE; P500048; PROTAMINE\_P1; 1.

DR PFAM; PF00260; protamine\_p1; 1.

KW CHROMOSOMAL PROTEIN; NUCLEOSOME CORE; SPERMATOGENESIS; DNA-BINDING;

TESTIS; DNA CONDENSATION; NUCLEAR PROTEIN.

FT INIT-MET 0 0

FT DISULFID 5 5 INTERCHAIN (WITH CYS-21) (BY SIMILARITY).

FT DISULFID 6 14 BY SIMILARITY.

FT DISULFID 21 21 INTERCHAIN (WITH CYS-5) (BY SIMILARITY).

FT DISULFID 36 36 INTERCHAIN (WITH CYS-36) (BY SIMILARITY).

FT DISULFID 37 47 BY SIMILARITY.

SQ SEQUENCE 50 AA; 6827 MW; D0D7B7C2 CRC32;

Query Match 57.6%; Score 34; DB 1; Length 50;

Best Local Similarity 66.7%; Pred. No. 8.1;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RTRCRFKRR 9

I:|:|:|:|:

Db 11 RSRRCRRRR 19

RESULT 15

HSP1\_RABIT

ID HSP1\_RABIT

AC P10119; STANDARD; PRT; 49 AA.

DT 01-MAR-1989 (REL. 10, CREATED)

DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)

DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)

DE SPERM PROTAMINE P1 (CYSTEINE-RICH PROTAMINE).

GN PRM1 OR PRM-1.

OS ORYCTOLAGUS CUNICULUS (RABBIT).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.

RN [1]

RP SEQUENCE.

RX MEDLINE; 89078594.

RA AMMER H., HENSCHEN A.;

RT "Primary structure of rabbit sperm protamine, the first protamine of  
 its type with an aberrant N-terminal.";

RL FEBS LETT. 242:111-116(1988).

CC -!- FUNCTION: PROTAMINES SUBSTITUTES FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.

CC -!- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE

CC DNA-HELIX (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: NUCLEAR.

CC -!- TISSUE SPECIFICITY: TESTIS.

DR PIR; S02007; S02007.

DR PROSITE; P500048; PROTAMINE\_P1; 1.

DR PFAM; PF00260; protamine\_p1; 1.

KW CHROMOSOMAL PROTEIN; NUCLEOSOME CORE; SPERMATOGENESIS; DNA-BINDING;

TESTIS; DNA CONDENSATION; NUCLEAR PROTEIN.

FT DISULFID 5 5 INTERCHAIN (WITH CYS-21) (BY SIMILARITY).

FT DISULFID 6 14 BY SIMILARITY.

FT DISULFID 21 21 INTERCHAIN (WITH CYS-5) (BY SIMILARITY).

FT DISULFID 37 37 INTERCHAIN (WITH CYS-37) (BY SIMILARITY).

FT DISULFID 38 46 BY SIMILARITY.

SQ SEQUENCE 49 AA; 6656 MW; 709F16CF CRC32;

Query Match 57.6%; Score 34; DB 1; Length 49;

Best Local Similarity 66.7%; Pred. No. 7.9;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RTRCRFKRR 9

I:|:|:|:|:

Db 11 RSRRCRRRR 19

Search completed: September 7, 1999, 23:50:20

Job time: 1959 sec

CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X15519; G61883; -  
 DR PIR; A34009; VHXRPD.  
 KW COAT PROTEIN; NONSTRUCTURAL PROTEIN.  
 SQ SEQUENCE 197 AA; 21458 MW; 71C39D31 CRC32;

Query Match 59.3%; Score 35; DB 1; Length 197;  
 Best Local Similarity 57.1%; Pred. No. 17;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 CREKRC 10  
 |::|||  
 Db 173 CKYKRC 179

RESULT 12  
 CR2\_HORVU STANDARD; PRT; 202 AA.  
 AC P23252;  
 DT 01-NOV-1991 (REL. 20, CREATED)  
 DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1991 (REL. 20, LAST ANNOTATION UPDATE)  
 DE COLD-REGULATED PROTEIN 2 (FRAGMENT).  
 OS HORDEUM VULGARE (BARLEY).  
 CC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 CC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;  
 CC POACEAE; HORDEUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. GEORGIE;  
 RA CATTIVELLI L., BARTELS D.;  
 RT "Molecular cloning and characterization of cold regulated genes in  
 RT barley.";  
 RL PLANT PHYSIOL. 93:1504-1510(1990).  
 CC -1- CONTAINS SEVERAL ARGININE RESIDUES IN CLOSE PROXIMITY WHICH  
 CC MAY BE INVOLVED IN PROTEIN RNA INTERACTIONS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M60733; G167028; ALT\_INIT.  
 DR PIR; B45512; B45512.  
 FT NON\_TER 1  
 FT DOMAIN 116 125 ARG-RICH.  
 SQ SEQUENCE 202 AA; 21753 MW; BC9A9416 CRC32;

Query Match 57.6%; Score 34; DB 1; Length 202;  
 Best Local Similarity 75.0%; Pred. No. 25;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TRCRFKRR 9  
 |::|||  
 Db 115 TRCRKR 122

RESULT 13  
 HSPL\_CAVPO STANDARD; PRT; 47 AA.  
 AC P35304;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE SPERM PROTAMINE P1 (CYSTEINE-RICH PROTAMINE).  
 GN PRM1 OR PRM-1.  
 OS CAVIA PORCELLUS (GUINEA PIG).

CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC RODENTIA; HYSTRICOGNATHI; CAVIIDAE; CAVIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JMI09;  
 RA QUERALT R., ADROER R., OLIVA R.;  
 RL SUBMITTED (DEC-1991) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- FUNCTION: PROTAMINES SUBSTITUTES FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE  
 CC DNA-HELIX.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: TESTIS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M83896; G607016; -  
 DR EMBL; Z11544; G49564; -  
 DR EMBL; Z11545; G49562; -  
 DR PIR; S29972; S29972.  
 DR PFAM; PF00260; PROTAMINE.P1; 1.  
 DR CHROMOSOMAL PROTEIN; NUCLEOSOME CORE; SPERMATOGENESIS; DNA-BINDING;  
 KW TESTIS; DNA CONDENSATION; NUCLEAR PROTEIN.  
 FT INIT\_MET 0  
 FT SEQUENCE 47 AA; 6553 MW; 49A375AE CRC32;

Query Match 57.6%; Score 34; DB 1; Length 47;  
 Best Local Similarity 66.7%; Pred. No. 7.7;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RTRCRFKRR 9  
 |::|||  
 Db 11 RSRRCRRR 19

RESULT 14  
 HSPL\_MOUSE STANDARD; PRT; 50 AA.  
 AC P02319;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE SPERM PROTAMINE P1 (CYSTEINE-RICH PROTAMINE).  
 GN PRM1 OR PRM-1.  
 OS MUS MUSCULUS (MOUSE).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 85199803.  
 RA KLEENE K.C., DISTEL R.J., HECHT N.B.;  
 RT "Nucleotide sequence of a cDNA clone encoding mouse protamine 1.";  
 RL BIOCHEMISTRY 24:719-722(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88193085.  
 RA JOHNSON P.A., PSCHON J.J., YELICK P.C., PALMITER R.D., HECHT N.B.;  
 RT "Sequence homologies in the mouse protamine 1 and 2 genes.";  
 RL BIOCHIM. BIOPHYS. ACTA 950:45-53(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 87260978.  
 RA PSCHON J.J., BEHRINGER R.R., BRINSTER R.L., PALMITER R.D.;  
 RT "Spermatid-specific expression of protamine 1 in transgenic mice.";

```

FT ZN-FING 87 111 C4-TYPE.
FT VASPLIC 409 419 CWPFRGAA -> S (IN SHORT FORM).
FT CONFLICT 171 171 K -> R (IN REF. 1).
FT CONFLICT 174 174 N -> S (IN REF. 1).
SQ SEQUENCE 465 AA; 51695 MW; A15DDEFF CRC32;

Query Match 59.3%; Score 35; DB 1; Length 465;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRFKRC 10
Db 87 CRFSROC 93
|||||

RESULT 9
OTCC_RHET STANDARD; PRT; 334 AA.
AC O31018;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE ORNITHINE CARBAMOYLTRANSFERASE, CATABOLIC (EC 2.1.3.3) (OTCASE).
GN ARCB.
OS RHIZOBIUM ETLI.
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;
OC RHIZOBIACEAE; RHIZOBIUM.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98053854.
RA D'HOOGHE I., VANDER WAUVE C., MICHELIS J., TRICOT C., DE WILDE P.,
VANDERLEIDEN J., STALON V.;
RT "The arginine deiminase pathway in Rhizobium etli: DNA sequence
analysis and functional study of the arcABC genes.";
RL J. BACTERIOL. 179:7403-7409(1997)
CC -!- CATALYTIC ACTIVITY: CARBAMOYL PHOSPHATE + ORNITHINE = CITRULLINE
+ ORTHOPHOSPHATE.
CC -!- PATHWAY: SECOND STEP IN ARGinine DEGRADATION VIA ARGinine
DEIMINASE.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: STRONG TO ALL OTHER OTCASES, ALSO HIGH TO ATCASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF025543; G2555171; .
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR PFAM; PF00185; OTCase; 1.
DR HSP; P08308; IORT.
KW TRANSFERASE; ARGinine METABOLISM.
SQ SEQUENCE 334 AA; 37817 MW; EAE32455 CRC32;

Query Match 59.3%; Score 35; DB 1; Length 334;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RTRCRFKRC 10
Db 59 RTRCAFEVAC 68
|||||

RESULT 10
RR14_EUGGR STANDARD; PRT; 100 AA.
ID RR14_EUGGR
AC P11538;
DT 01-OCT-1989 (REL. 12, CREATED)

```

```

DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S14.
GN RPS14.
OS EUGLENA GRACILIS.
OC CHLOROPLAST.
OC EURARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z;
RX MEDLINE; 89315233.
RA NICKOLOFF J.A., CHRISTOPHER D.A., DRAGER R.G., HALLICK R.B.;
RT "Nucleotide sequence of the Euglena gracilis chloroplast genes for
isooleucine, phenylalanine and cysteine transfer RNAs and ribosomal
protein S14.";
RT NUCLEIC ACIDS RES. 17:4882-4882(1989).
RL -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; Z11874; G14377; .
DR EMBL; X15240; G11529; .
DR EMBL; X70810; G415766; .
DR PIR; S04701; R3EG14.
DR PIR; S34531; S34531.
DR PROSITE; PS00527; RIBOSOMAL_S14; 1.
DR PFAM; PF00253; S14; 1.
DR MENDEL; 4766; EUGgr;rps14.1.
KW RIBOSOMAL PROTEIN; CHLOROPLAST.
SQ SEQUENCE 100 AA; 12014 MW; BB520AD5 CRC32;

Query Match 59.3%; Score 35; DB 1; Length 100;
Best Local Similarity 62.5%; Pred. No. 97;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRFKRC 10
Db 56 RCRLNRC 63
|||||

RESULT 11
VS11_ROT5 STANDARD; PRT; 197 AA.
AC P19715;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE MINOR OUTER CAPSID PROTEIN (NS26) (NONSTRUCTURAL PROTEIN VP9).
GN S11.
OS PORCINE ROTAVIRUS (SEROTYPE 5 / STRAIN OSU).
OC VIRUSES; DSRNA VIRUSES; REOVIRIDAE; ROTAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89366681.
RA GONZALEZ S.A., BURRONE O.R.;
RT "Porcine OSU rotavirus segment II sequence shows common features with
the viral gene of human origin.";
RT NUCLEIC ACIDS RES. 17:6402-6402(1989).
RL -!- SIMILARITY: TO OTHER ROTAVIRUSES MINOR OUTER CAPSID PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

DR HSP: P19793; 2NLL.  
 DR TRANSFAC; T00373; -.  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 ZINC-FINGER; LIVER; ALTERNATIVE SPLICING.  
 FT DNA\_BIND 51 116 C4-TYPE ZINC FINGERS (TWO).  
 FT 2N\_FING 51 71 C4-TYPE.  
 FT 2N\_FING 87 111 C4-TYPE.  
 FT VARSPLIC 409 419 CEMPRGQAA -> S (IN SHORT FORM).  
 SQ SEQUENCE 465 AA; 51754 MW; 1FD7F232 CRC32;

Query Match 59.3%; Score 35; DB 1; Length 465;  
 Best Local Similarity 71.4%; Pred. No. 33;  
 Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 4 CRFKRRC 10  
 ||| |:  
 Db 87 CRFSROC 93

RESULT 7  
 HNF4\_MOUSE  
 ID HNF4\_MOUSE STANDARD; PRT; 465 AA.  
 AC P49698;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE HEPATOCYTE NUCLEAR FACTOR 4 (TRANSCRIPTION FACTOR HNF-4)  
 DE (TRANSCRIPTION FACTOR 14).  
 GN TCF14 OR HNF4 OR HNF-4.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B6/CBA; TISSUE=LIVER;  
 RX MEDLINE; 93092794.  
 RA HATA S., INOUE T., KOSUGA K., NAKASHIMA T., TSUKAMOTO T., OSUMI T.;  
 RT "Identification of two splice isoforms of mRNA for mouse hepatocyte  
 nuclear factor 4 (HNF-4).";  
 RL BIOCHIM. BIOPHYS. ACTA 1260:55-61(1995).  
 CC -!- FUNCTION: TRANSCRIPTIONALLY CONTROLLED TRANSCRIPTION FACTOR. BINDS  
 TO DNA SITES REQUIRED FOR THE TRANSCRIPTION OF ALPHA 1-  
 ANTITRYPsin, APOLIPOPROTEIN CIII AND TRANSTHYRETIN GENES. MAY BE  
 ESSENTIAL FOR DEVELOPMENT OF THE LIVER, KIDNEY AND INTESTINE.  
 CC -!- SUBUNIT: HOMODIMERIZATION IS REQUIRED FOR HNF-4 TO BIND TO ITS  
 RECOGNITION SITE (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY ALTERNATIVE  
 SPLICING.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D29015; G511934; -.  
 DR MGD; MGI:109128; HNF4.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSP; P19793; 2NLL.  
 DR TRANSFAC; T02423; -.  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 ZINC-FINGER; LIVER; ALTERNATIVE SPLICING.  
 FT DNA\_BIND 51 116 C4-TYPE ZINC FINGERS (TWO).  
 FT 2N\_FING 51 71 C4-TYPE.  
 FT 2N\_FING 87 111 C4-TYPE.  
 FT VARSPLIC 409 419 CEMPRGQAA -> S (IN SHORT FORM).

SQ SEQUENCE 465 AA; 51755 MW; 097865A9 CRC32;

Query Match 59.3%; Score 35; DB 1; Length 465;  
 Best Local Similarity 71.4%; Pred. No. 33;  
 Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 4 CRFKRRC 10  
 ||| |:  
 Db 87 CRFSROC 93

RESULT 8  
 HNF4\_RAT  
 ID HNF4\_RAT STANDARD; PRT; 465 AA.  
 AC P22449;  
 DT 01-AUG-1991 (REL. 19, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HEPATOCYTE NUCLEAR FACTOR 4 (TRANSCRIPTION FACTOR HNF-4)  
 DE (TRANSCRIPTION FACTOR 14).  
 GN TCF14 OR HNF4 OR HNF-4.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=LIVER;  
 RX MEDLINE; 91122637.  
 RA SLADEK F.M., ZHONG W., LAI E., DARNELL J.E. JR.;  
 RT "Liver-enriched transcription factor HNF-4 is a novel member of the  
 steroid hormone receptor superfamily.";  
 RL GENES DEV. 4:2353-2365(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC STRAIN=Wistar; TISSUE=LIVER;  
 RX MEDLINE; 92305063.  
 RA HATA S., TSUKAMOTO T., OSUMI T.;  
 RT "A novel isoform of rat hepatocyte nuclear factor 4 (HNF-4).";  
 RL BIOCHIM. BIOPHYS. ACTA 1131:211-213(1992).  
 CC -!- FUNCTION: TRANSCRIPTIONALLY CONTROLLED TRANSCRIPTION FACTOR. BINDS  
 TO DNA SITES REQUIRED FOR THE TRANSCRIPTION OF ALPHA 1-  
 ANTITRYPsin, APOLIPOPROTEIN CIII AND TRANSTHYRETIN GENES. MAY BE  
 ESSENTIAL FOR DEVELOPMENT OF THE LIVER, KIDNEY AND INTESTINE.  
 CC -!- SUBUNIT: HOMODIMERIZATION IS REQUIRED FOR HNF-4 TO BIND TO ITS  
 RECOGNITION SITE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- TISSUE SPECIFICITY: LIVER, KIDNEY AND INTESTINE.  
 CC -!- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY ALTERNATIVE  
 SPLICING.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D10554; G220773; -.  
 DR EMBL; X57133; G56372; -.  
 DR PIR; A36471; A36471.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSP; P19793; 2NLL.  
 DR TRANSFAC; T02422; -.  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 ZINC-FINGER; LIVER; ALTERNATIVE SPLICING.  
 FT MOD\_RES 21 21 BLOCKED.  
 FT DNA\_BIND 51 116 C4-TYPE ZINC FINGERS (TWO).  
 FT 2N\_FING 51 71 C4-TYPE.

RT in Korea.":  
RL NUCLEIC ACIDS RES. 20:5474-5474(1992).  
CC -!- FUNCTION: PROBABLE POLYMERASE.  
CC -!- READTHROUGH OF THE TERMINATOR CODON UAG OCCURS BETWEEN CODONS FOR  
CC GLN-1116 AND GLN-1117.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X68110; -; NOT-ANNOTATED\_CDS.  
DR PFAM: PF00978; RNA\_dep\_RNAPol2; 1.  
DR TRANSFERASE; RNA-DIRECTED RNA POLYMERASE; HELICASE; ATP-BINDING.  
KW CHAIN 1 1615 180 KD PROTEIN.  
FT CHAIN 1 1116 130 KD PROTEIN.  
FT NP\_BIND 833 840 ATP (POTENTIAL).  
SQ SEQUENCE 1615 AA; 183392 MW; DB10FEE CRC32;  
-----  
Query Match 61.0%; Score 36; DB 1; Length 1615;  
Best Local Similarity 85.7%; Pred. No. 63;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
-----  
QY 2 TRCRFKR 8  
Db 896 TRCQFKR 902  
-----  
RESULT 5  
DEF1\_RAT  
ID DEF1\_RAT STANDARD; PRT; 94 AA.  
AC Q62716;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE NEUTROPHIL ANTIBIOTIC PEPTIDE NP-1 PRECURSOR.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 63-68 AND 86-91.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=BONE MARROW;  
RX MEDLINE; 96025910.  
RA YOUNT N.Y., WANG M.-S.C., YUAN J., BANAIIE N., OUELLETTE A.J.,  
RA SELSTED M.E.;  
RT "Rat neutrophil defensins. Precursor structures and expression during  
RT neutrophilic myelopoiesis."  
RL J. IMMUNOL. 155:4476-4484(1995).  
RN [2]  
RP SEQUENCE OF 63-94.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=PERITONEAL NEUTROPHILS;  
RX MEDLINE; 89277517.  
RA EISENHAUER P.B., HARWIG S.S.-L., SZKLAREK D., GANZ T., SELSTED M.E.,  
RA LEHRER R.I.;  
RT "Purification and antimicrobial properties of three defensins from  
RT rat neutrophils."  
RL INFECT. IMMUN. 57:2021-2027(1989).  
CC -!- FUNCTION: ACTIVE IN VITRO AGAINST S.AUREUS, FUNGII, GRAM-POSITIVE  
CC AND GRAM-NEGATIVE BACTERIA AND TO A LESSER EXTENT AGAINST  
CC AN ENVELOPED VIRUS.  
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BONE MARROW AND TO A  
CC MUCH LESSER EXTENT IN SMALL INTESTINE.  
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF CORTICOSTATIN/DEFENSIN  
CC PEPTIDES.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U16686; G1041811; -;  
DR PROSITE; PS00269; DEFENSIN; 1.  
DR PFAM; PF00323; defensins; 1.  
DR PFAM; PF00879; Defensin-Propep; 1.  
KW DEFENSIN; SIGNAL; ANTIBIOTIC; FUNGICIDE.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 62 POTENTIAL.  
FT CHAIN 63 94 NEUTROPHIL ANTIBIOTIC PEPTIDE NP-1.  
FT DISULFID 65 93 BY SIMILARITY.  
FT DISULFID 67 82 BY SIMILARITY.  
FT DISULFID 72 92 BY SIMILARITY.  
SQ SEQUENCE 94 AA; 10371 MW; BE075BC7 CRC32;  
-----  
Query Match 59.3%; Score 35; DB 1; Length 94;  
Best Local Similarity 66.7%; Pred. No. 9.2;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
-----  
QY 1 RTRCRFKRR 9  
Db 69 RTRCGFRER 77  
-----  
RESULT 6  
HNF4\_HUMAN  
ID HNF4\_HUMAN STANDARD; PRT; 465 AA.  
AC P41235;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HEPATOCYTE NUCLEAR FACTOR 4 (TRANSCRIPTION FACTOR HNF-4)  
DE (TRANSCRIPTION FACTOR 14).  
GN TCF14 OR HNF4.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE; 95011627.  
RA CHARTIER F.L., BOSSU J.-P., LAUDET V., FRUCHART J.-C., LAINE B.;  
RT "Cloning and sequencing of cDNAs encoding the human hepatocyte  
RT nuclear factor 4 indicate the presence of two isoforms in human  
RT liver."  
RL GENE 147:269-272(1994).  
CC -!- FUNCTION: TRANSCRIPTIONALLY CONTROLLED TRANSCRIPTION FACTOR. BINDS  
CC TO DNA SITES REQUIRED FOR THE TRANSCRIPTION OF ALPHA 1-  
CC ANTITRYPSIN, APOLOPROTEIN CIII AND TRANSFERRIN GENES. MAY BE  
CC ESSENTIAL FOR DEVELOPMENT OF THE LIVER, KIDNEY AND INTESTINE.  
CC -!- SUBUNIT: HOMODIMERIZATION IS REQUIRED FOR HNF-4 TO BIND TO ITS  
CC RECOGNITION SITE.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY ALTERNATIVE  
CC SPLICING.  
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X76930; G575253; -;  
DR MIN; 600281; -;  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR PFAM; PF00104; hormone\_rec; 1.  
DR PFAM; PF00105; zf-C4; 1.



Db 78 TRCEDRDRTRRC 90

```

RESULT 2
POCR_SALTY STANDARD; PRT; 303 AA.
ID POCR_SALTY
AC Q05587;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE REGULATORY PROTEIN POCR.
GN POCR.
OS SALMONELLA TYPHIMURIUM.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC SALMONELLA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-LT2;
RC STRAIN-LT2;
RA CHEN P., ANDERSON D.I., ROTH J.R.;
RT "The control region of the pdu/cob regulon in Salmonella typhimurium.";
RL J. BACTERIOL. 176:5474-5482(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-LT2;
RC STRAIN-LT2;
RA BOBIK T.A., XU Y., JETER R.M., OTTO K.E., ROTH J.R.;
RT "Propanediol utilization genes (pdu) of Salmonella typhimurium: three genes for the propanediol dehydratase.";
RL J. BACTERIOL. 179:6633-6639(1997).
RN [3]
RP SEQUENCE OF 75-303 FROM N.A.
RX STRAIN-LT2;
RC STRAIN-LT2;
RA ROTH J.R., LAWRENCE J.G., RUBENFIELD M., KIEFFER-HIGGINS S., CHURCH G.M.;
RT "Characterization of the cobalamin (vitamin B12) biosynthetic genes of Salmonella typhimurium.";
RL J. BACTERIOL. 175:3303-3316(1993).
CC -|- FUNCTION: POSITIVE REGULATORY PROTEIN OF PDU AND COB OPERONS.
CC -|- PATHWAY: COBALAMIN BIOSYNTHESIS.
CC -|- PATHWAY: GLYCEROL AND PROPANEDIOL UTILIZATION.
CC -|- SIMILARITY: BELONGS TO THE ARAC/XLYS FAMILY OF TRANSCRIPTIONAL REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF026270; G2587036;
DR EMBL; L12006; G154419;
DR STYGENE; SG10307; POCR.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
DR PFAM; PF00165; HTH_2; 1.
DR COBALAMIN BIOSYNTHESIS; TRANSCRIPTION REGULATION; ACTIVATOR;
KW DNA-BINDING; GLYCEROL METABOLISM.
FT DNA_BIND 211 230 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 303 AA; 34414 MW; 18D62808 CRC32;

Query Match 61.0%; Score 36; DB 1; Length 303;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TRCRFKRC 10
DB 58 TRCRMSDRC 66

```

```

RESULT 3
RRPO_TMVR STANDARD; PRT; 1615 AA.
ID RRPO_TMVR
AC P03586;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE PUTATIVE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (180 KD PROTEIN) [CONTAINS: 130 KD PROTEIN; RNA HELICASE].
OS TOBACCO MOSAIC VIRUS (VULGARE) (TMV).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; TOBAMOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 83299880.
RA GOELET P., LOMONOSOFF G.P., BUTLER P.J.G., AKAM M.E., GAIT M.J., KARN J.;
RT "Nucleotide sequence of tobacco mosaic virus RNA.";
RL PROC. NATL. ACAD. SCI. U.S.A. 79:5818-5822(1982).
CC -|- FUNCTION: PROBABLE POLYMERASE.
CC -|- READTHROUGH OF THE TERMINATOR CODON UAG OCCURS BETWEEN CODONS FOR GLN-1116 AND GLN-1117.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01408; G663264; ALT.FRAME.
DR EMBL; V01409; -: NOT_ANNOTATED_CDS.
DR PIR; A04194; WMTM18.
DR PFAM; PF00978; RNA_DIRECTED_RNA_POLYMERASE; ATP-BINDING.
KW TRANSFERASE; RNA-DIRECTED RNA POLYMERASE; HELICASE;
FT CHAIN 1 1615 180 KD PROTEIN.
FT CHAIN 1 1116 130 KD PROTEIN.
FT NP_BIND 833 840 ATP (POTENTIAL).
FT VARIANT 9 9 T -> S.
FT VARIANT 14 14 D -> E.
FT VARIANT 21 21 S -> T.
FT VARIANT 37 37 E -> D.
FT VARIANT 638 638 E -> A.
SQ SEQUENCE 1615 AA; 183355 MW; 9C23E037 CRC32;

Query Match 61.0%; Score 36; DB 1; Length 1615;
Best Local Similarity 85.7%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TRCRFKR 8
DB 896 TRCQFKR 902

RESULT 4
RRPO_TMVR STANDARD; PRT; 1615 AA.
ID RRPO_TMVR
AC P30738;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE PUTATIVE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (180 KD PROTEIN) [CONTAINS: 130 KD PROTEIN; RNA HELICASE].
OS TOBACCO MOSAIC VIRUS (STRAIN KOREAN) (TMV).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; TOBAMOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93065219.
RA KOB K.H., SONG E.K., LEE S.Y., PARK Y.I., PARK W.M.;
RT "Nucleotide sequence of cDNA of the tobacco mosaic virus RNA isolated

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:19 ; Search time 71.87 Seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-124-280A-17

Perfect score: 59

Sequence: 1 RTRCRFKRRC 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	61.0	381	DUS6_RAT	Q64346 rattus norv
2	36	61.0	303	POCR_SALTY	Q05587 salmonella
3	36	61.0	1615	RRPO_TNVR	P03586 tobacco mos
4	36	61.0	1615	RRPO_TNVR	P30738 tobacco mos
5	35	59.3	94	DEF1_RAT	Q62716 rattus norv
6	35	59.3	465	HNF4_HUMAN	P41235 homo sapien
7	35	59.3	465	HNF4_MOUSE	P49698 mus muscucu
8	35	59.3	465	HNF4_RAT	P22449 rattus norv
9	35	59.3	334	OTCC_RHET	O31018 rhizobium e
10	35	59.3	100	RR14_EUGGR	P11538 euglena gra
11	35	59.3	197	VS11_ROTTP5	P19175 porcine rot
12	34	57.6	202	CR2_HORVU	P23252 hordeum vul
13	34	57.6	47	HSP1_CAVPO	P35304 cavia porce
14	34	57.6	50	HSP1_MOUSE	P02319 mus muscucu
15	34	57.6	49	HSP1_RABIT	P10119 oryctolagus
16	34	57.6	50	HSP1_RAT	P10118 rattus norv
17	34	57.6	50	HSP1_SHEEP	P04102 ovis aries
18	34	57.6	887	OBP_HSVB	P28947 equine herp
19	34	57.6	976	SN21_HUMAN	P28370 homo sapien
20	34	57.6	296	UL07_HSV11	P10191 herpes simp
21	34	57.6	193	YAK7_YEAST	P39553 saccharomyc
22	34	57.6	247	YJ35_MYCLE	P54075 mycobacteri
23	33.5	56.8	450	TH12_YEAST	P38141 saccharomyc
24	33	55.9	235	BAG6_CHITE	P08726 chironomus
25	33	55.9	183	DERM_BOVIN	P19427 bos taurus
26	33	55.9	201	DERM_HUMAN	Q07507 homo sapien
27	33	55.9	44	DERM_PIG	P45846 sus scrofa
28	33	55.9	666	HNF4_DROME	P49866 drosophila
29	33	55.9	233	H0XL_ALCEU	P22318 alcaligenes
30	33	55.9	66	HSP1_ISOMA	P42136 isodon mac
31	33	55.9	46	HSP1_ORCOR	P24713 orcinus orc
32	33	55.9	68	HSP1_PERGU	P42147 perameles g
33	33	55.9	49	HSP1_PIG	P04101 sus scrofa
34	33	55.9	1094	YB00_YEAST	P38114 saccharomyc
35	32.5	55.1	549	TEGU_HCMV	P07387 human cytom
36	32	54.2	564	AGGL_RICCO	P06750 ricinus com
37	32	54.2	1394	CNG4_BOVIN	Q28181 bos taurus
38	32	54.2	51	DIS1_ERIMA	P22826 eristocophi
39	32	54.2	347	GBX2_HUMAN	P52951 homo sapien
40	32	54.2	1520	GLTB_BACSU	P39812 bacillus su
41	32	54.2	284	HAT4_ARATH	Q05466 arabidopsis
42	32	54.2	51	HSP1_ALOSE	P35302 alouatta se
43	32	54.2	50	HSP1_BOVIN	P02318 bos taurus

44 32 54.2 49 1 HSP1\_HORSE P15341 equus cabal  
45 32 54.2 104 1 HSP2\_CALJA Q28337 callithrix

#### ALIGNMENTS

```

RESULT 1
DUS6_RAT STANDARD; PRT; 381 AA.
AC Q64346;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 6 (EC 3.1.3.48) (EC 3.1.3.16)
DE (MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE 3) (MAP KINASE
DE PHOSPHATASE 3) (MKP-3).
DE DUSP6 OR MKP3.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-NEURON;
RX MEDLINE; 96224012.
RA MUDA M., BOSCHERT U., DICKINSON R., MARTINOU J.C., MARTINOU I.,
RA CAMPS M., SCHLEGEL W., ARKINSTALL S.;
RT "MKP-3", a novel cytosolic protein-tyrosine phosphatase that
RT exemplifies a new class of mitogen-activated protein kinase
RT phosphatases."
RT J. BIOL. CHEM. 271:4319-4326(1996).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96216487.
RA MOUREY R.J., VEGA Q.C., CAMPBELL J.S., WENDEROTH M.P., HAUSCHKA S.D.,
RA KREBS E.G., DIXON J.E.;
RT "A novel cytoplasmic dual specificity protein tyrosine phosphatase
RT implicated in muscle and neuronal differentiation."
RL J. BIOL. CHEM. 271:3795-3802(1996).
CC -!- FUNCTION: INACTIVATES MAP KINASES. HAS A SPECIFICITY FOR THE ERK
CC FAMILY. IMPLICATED IN MUSCLE AND NEURONAL DIFFERENTIATION.
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X94185; E218443; -
DR EMBL; U42627; G1185552; -
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR PFAM; PF00782; DSPC; 1.
DR HSPSP; P51452; 1VHR.
DR KW HYDROLASE.
FT ACT_SITE. 293 293 BY SIMILARITY.
SQ SEQUENCE 381 AA; 42318 MW; B036E72B CRC32;

```

Query Match 61.0%; Score 36; DB 1; Length 381;  
Best Local Similarity 61.5%; Pred No. 19;  
Matches 8; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

OY 2 TRC----RFKRC 10  
||| |||||

```

RESULT 12
JC6095
hepatocyte nuclear factor 4 gamma chain - human
C:Species: Homo sapiens (man)
C:Date: 13-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 12-Sep-1997
C:Accession: JC6095
R:Drawes, T.; Senkel, S.; Holey, B.; Ryffel, G.O.
Mol. Cell. Biol. 16, 925-931, 1996
A:Title: Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and different
A:Reference number: JC6095; MUID:96182096
A:Accession: JC6095
A:Molecule type: mRNA
A:Residues: 1-774 <KRI>
A:Cross-references: EMBL:249826; NID:gl1217962; PID:el83832; PID:gl1217963
A:Experimental source: kidney
A:Comment: This factor is an orphan member of the nuclear receptor superfamily with a zi
sis, and plays an important role in early embryogenesis.
C:Genetics:
A:Gene: hnf4 gamma
A:Map position: 8
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: DNA binding; kidney; transcription factor; zinc finger
F:376-614/Domain: erba transforming protein homology <ERBA>

Query Match 59.3%; Score 35; DB 2; Length 774;
Best Local Similarity 71.4%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 4 CRFKRRC 10
Db 414 CRFSROC 420
||| | |
| | | | |

RESULT 13
JC4936
hepatocyte nuclear factor 4A - human
C:Species: Homo sapiens (man)
C:Date: 22-Oct-1996 #sequence_revision 01-Nov-1996 #text_change 31-Oct-1997
C:Accession: JC4936
R:Kritsis, A.A.; Argyrokastritis, A.; Moschonas, N.K.; Power, S.; Katrakilli, N.; Zannis,
Gene 173, 275-280, 1996
A:Title: Isolation and characterization of a third isoform of human hepatocyte nuclear
A:Reference number: JC4936; MUID:97082982
A:Accession: JC4936
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-455 <KRI>
A:Cross-references: EMBL:X87870; NID:gl1595751; PID:el84013; PID:gl1595752
A:Experimental source: liver
A:Note: nucleic acid sequence is not complete in this paper
C:Comment: This protein is one of the positive regulators of liver-specific genes.
C:Genetics:
A:Gene: hhnf-4A
A:Map position: 20
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: zinc finger
F:49-288/Domain: erba transforming protein homology <ERBA>

Query Match 59.3%; Score 35; DB 2; Length 455;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 4 CRFKRRC 10
Db 87 CRFSROC 93
||| | |
| | | | |

RESULT 14
JC4937
hepatocyte nuclear factor 4B - human
C:Species: Homo sapiens (man)

```

```

C:Date: 22-Oct-1996 #sequence_revision 01-Nov-1996 #text_change 31-Oct-1997
C:Accession: JC4937
R:Kritsis, A.A.; Argyrokastritis, A.; Moschonas, N.K.; Power, S.; Katrakilli, N.; Zanni
Gene 173, 275-280, 1996
A:Title: Isolation and characterization of a third isoform of human hepatocyte nuclea
A:Reference number: JC4936; MUID:97082982
A:Accession: JC4937
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-465 <KRI>
A:Cross-references: EMBL:X87871; NID:gl1595755; PID:el84014; PID:gl1595756
A:Experimental source: liver
A:Note: nucleic acid sequence is not complete in this paper
C:Comment: This protein is one of the positive regulators of liver-specific genes.
C:Genetics:
A:Gene: hhnf-4B
A:Map position: 20
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: zinc finger
F:49-288/Domain: erba transforming protein homology <ERBA>

```

```

Query Match 59.3%; Score 35; DB 2; Length 465;
Best Local Similarity 71.4%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

```

```

QY 4 CRFKRRC 10
Db 87 CRFSROC 93
||| | |
| | | | |

```

```

RESULT 15
JC4938
hepatocyte nuclear factor 4C - human
C:Species: Homo sapiens (man)
C:Date: 22-Oct-1996 #sequence_revision 01-Nov-1996 #text_change 31-Oct-1997
C:Accession: JC4938
R:Kritsis, A.A.; Argyrokastritis, A.; Moschonas, N.K.; Power, S.; Katrakilli, N.; Zanni
Gene 173, 275-280, 1996
A:Title: Isolation and characterization of a third isoform of human hepatocyte nuclea
A:Reference number: JC4936; MUID:97082982
A:Accession: JC4938
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-408 <KRI>
A:Cross-references: EMBL:X87872; NID:gl1595753; PID:el84046; PID:gl1595754
A:Experimental source: liver
A:Note: nucleic acid sequence not complete
C:Comment: This protein is one of the positive regulators of liver-specific genes.
C:Genetics:
A:Gene: hhnf-4C
A:Map position: 20
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: zinc finger
F:49-288/Domain: erba transforming protein homology <ERBA>

```

```

Query Match 59.3%; Score 35; DB 2; Length 408;
Best Local Similarity 71.4%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

```

```

QY 4 CRFKRRC 10
Db 87 CRFSROC 93
||| | |
| | | | |

```

```

Search completed: September 7, 1999, 23:06:31
Job time: 2479 sec

```

Gene 147, 269-272, 1994  
A:Title: Cloning and sequencing of cDNAs encoding the human hepatocyte nuclear factor 4  
A:Reference number: JC4009; MUID:95011627  
A:Accession: JC4009  
A:Molecule type: mRNA  
A:Residues: 1-463 <CHA>  
A:Cross-references: EMBL:X76930  
A:Experimental source: liver  
C:Genetics:  
A:Gene: GDB:TCF14; HNF4  
A:Cross-references: GDB:393281; OMIM:600281  
A:Map position: 20q12-20q13.1  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: transcription factor; zinc finger  
F:47-286/Domain: erba transforming protein homology <ERBA>

Query Match 59.3%; Score 35; DB 2; Length 463;  
Best Local Similarity 71.4%; Pred. No. 49;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRFKRRC 10  
||| |||  
Db 85 CRFSRQC 91

RESULT 8  
S52074  
hepatocyte nuclear factor 4 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 31-Oct-1997  
C:Accession: S52074  
R:Hata, S.; Inoue, T.; Kosuga, K.; Nakashima, T.; Tsukamoto, T.; Osumi, T.  
Biochim. Biophys. Acta 1260, 55-61, 1995  
A:Title: Identification of two splice isoforms of mRNA for mouse hepatocyte nuclear factor 4  
A:Reference number: S52074; MUID:95092794  
A:Accession: S52074  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-465 <HAT>  
A:Cross-references: EMBL:D29015; NID:q469216; PID:d1006655; PID:g511934  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: zinc finger  
F:49-288/Domain: erba transforming protein homology <ERBA>

Query Match 59.3%; Score 35; DB 2; Length 465;  
Best Local Similarity 71.4%; Pred. No. 49;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRFKRRC 10  
||| |||  
Db 87 CRFSRQC 93

RESULT 9  
S23502  
hepatocyte nuclear factor 4 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 31-Oct-1997  
C:Accession: S23502  
R:Hata, S.; Tsukamoto, T.; Osumi, T.  
Biochim. Biophys. Acta 1131, 211-213, 1992  
A:Title: A novel isoform of rat hepatocyte nuclear factor 4 (HNF-4).  
A:Reference number: S23502; MUID:92305063  
A:Accession: S23502  
A:Molecule type: mRNA  
A:Residues: 1-465 <HAT>  
A:Cross-references: EMBL:D10554; NID:g220772; PID:d1001884; PID:g220773  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: zinc finger  
F:49-288/Domain: erba transforming protein homology <ERBA>

Query Match 59.3%; Score 35; DB 2; Length 465;  
Best Local Similarity 71.4%; Pred. No. 49;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRFKRRC 10  
||| |||  
Db 87 CRFSRQC 93

RESULT 10  
A36471  
transcription factor HNF-4 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 19-Apr-1991 #sequence\_revision 19-Apr-1991 #text\_change 12-Sep-1997  
C:Accession: A36471  
R:Sladek, F.M.; Zhong, W.; Lai, E.; Darnell Jr., J.E.  
Genes Dev. 4, 2353-2365, 1990  
A:Title: Liver-enriched transcription factor HNF-4 is a novel member of the steroid h  
A:Reference number: A36471; MUID:91122637  
A:Accession: A36471  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-455 <SLA>  
A:Cross-references: GB:X57133; NID:g56371; PID:g56372  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: DNA binding; transcription regulation; zinc finger  
F:49-288/Domain: erba transforming protein homology <ERBA>

Query Match 59.3%; Score 35; DB 2; Length 455;  
Best Local Similarity 71.4%; Pred. No. 48;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRFKRRC 10  
||| |||  
Db 87 CRFSRQC 93

RESULT 11  
JC6096  
hepatocyte nuclear factor 4 alpha2 chain - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 12-Sep-1997  
C:Accession: JC6096  
R:Drewes, T.; Senkel, S.; Holewa, B.; Ryffel, G.U.  
Mol. Cell. Biol. 16, 925-931, 1996  
A:Title: Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and diffe  
A:Reference number: JC6095; MUID:96182096  
A:Accession: JC6096  
A:Molecule type: mRNA  
A:Residues: 1-504 <DRE>  
A:Cross-references: EMBL:Z49825; NID:gl217960; PID:el83831; PID:gl217961  
A:Experimental source: kidney  
C:Comment: This factor is a tissue-specific transcription factor, which belongs to or  
and plays an important role in early embryogenesis.  
C:Genetics:  
A:Gene: hnfa alpha2  
A:Map position: 20  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: DNA binding; kidney; transcription factor; zinc finger  
F:88-327/Domain: erba transforming protein homology <ERBA>

Query Match 59.3%; Score 35; DB 2; Length 504;  
Best Local Similarity 71.4%; Pred. No. 52;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRFKRRC 10  
||| |||  
Db 126 CRFSRQC 132

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:30 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-17  
Perfect score: 59  
Sequence: 1 RTRCRFRRC 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	61.0	1615	1 WMTM18	183.3K protein - t
2	36	61.0	1116	2 S26358	hypothetical prote
3	36	61.0	1538	2 H70846	hypothetical glyci
4	35	59.3	100	1 R3EG14	ribosomal protein
5	35	59.3	197	1 VHXRPU	minor outer capsid
6	35	59.3	32	2 D61014	defensin NP-1 - ra
7	35	59.3	463	2 J24009	hepatocyte nuclear
8	35	59.3	465	2 S22074	hepatocyte nuclear
9	35	59.3	465	2 S23502	hepatocyte nuclear
10	35	59.3	455	2 A36471	transcription fact
11	35	59.3	504	2 J26096	hepatocyte nuclear
12	35	59.3	774	2 J26095	hepatocyte nuclear
13	35	59.3	455	2 J24936	hepatocyte nuclear
14	35	59.3	465	2 J24937	hepatocyte nuclear
15	35	59.3	408	2 J24938	hepatocyte nuclear
16	35	59.3	1118	2 A48292	mucin, tracheobron
17	34	57.6	50	1 HSSH	sperm histone - sh
18	34	57.6	51	1 HSMSS1	protamine - mouse
19	34	57.6	296	1 WMBEX7	UL7 protein - huma
20	34	57.6	887	1 WBEES	gene 53 protein -
21	34	57.6	48	2 S29973	protamine 1 - guin
22	34	57.6	49	2 S02007	protamine 1 - rabb
23	34	57.6	51	2 S03997	protamine 1 - rat
24	34	57.6	535	2 S18606	phosphoenolpyruv
25	34	57.6	151	2 A64503	conserved hypothet
26	34	57.6	247	2 S72859	hypothetical prote
27	34	57.6	435	2 C70509	hypothetical prote
28	34	57.6	313	2 A69471	dipeptide ABC tran
29	34	57.6	202	2 B45512	cold-regulated pro
30	34	57.6	976	2 S35457	SNF2 protein homol
31	33.5	56.8	450	2 S46117	probable regulator
32	33	55.9	50	1 HSPG	sperm histone - pi
33	33	55.9	234	2 B35385	hydrogen dehydroge
34	33	55.9	47	2 S22828	protamine - killer
35	33	55.9	50	2 S21672	protamine 1 - pig
36	33	55.9	666	2 S36218	hepatocyte nuclear
37	33	55.9	1094	2 S46021	probable regulator
38	33	55.9	508	2 A45477	CAMP response elem
39	33	55.9	369	2 I78877	CAMP responsive el

ALIGNMENTS

RESULT 1

WMTM18

183.3K protein - tobacco mosaic virus

C:Species: tobacco mosaic virus, TMV

C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 15-Oct-1994

C:Accession: A04194

R:Goelet, P.; Lomonosoff, G.P.; Butler, P.J.G.; Akam, M.E.; Gait, M.J.; Karn, J.

Proc. Natl. Acad. Sci. U.S.A. 79, 5818-5822, 1982

A:Title: Nucleotide sequence of tobacco mosaic virus RNA.

A:Reference number: A93926; MUID:8329980

A:Accession: A04194

A:Molecule type: genomic RNA

A:Residues: 1-1615 <GOE>

A:Note: the authors translated the codon AUU for residue 1302 as Tyr

A:Note: this sequence contains proteins 125.9K (residues 1-1116), 87.4K (336-1116), 8

A:Note: readthrough of the terminator codon UAG between codons for Gln-1116 and Gln-1

A:Note: variant forms are found with 9-Ser, 14-Glu, 21-Thr, 37-Asp, and 683-Ala

C:Superfamily: cucumber mosaic virus RNA 1 protein

Query Match 61.0%; Score 36; DB 1; Length 1615;

Best Local Similarity 85.7%; Pred. No. 87;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 TRCRFKR 8

Db 896 TRCQPKR 902

RESULT 2

S26358

hypothetical protein, 126K - tobacco mosaic virus

C:Species: tobacco mosaic virus, TMV

C:Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 10-Nov-1995

C:Accession: S26358

R:Koh, H.K.; Song, E.K.; Lee, S.Y.; Park, Y.I.; Park, W.M.

Nucleic Acids Res. 20, 5474, 1992

A:Title: Nucleotide sequence of cDNA of the tobacco mosaic virus RNA isolated in Kore

A:Reference number: S26358

A:Accession: S26358

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-1116 <KOH>

A:Cross-references: EMBL:X68110

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992

Query Match 61.0%; Score 36; DB 2; Length 1116;

Best Local Similarity 85.7%; Pred. No. 65;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 TRCRFKR 8

Db 896 TRCQPKR 902

RESULT 3

H70846

hypothetical glycine-rich protein Rv3345c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 24-Jul-1998

C:Accession: H70846

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
 A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A: Reference number: A70500; MUID: 98295987  
 A: Accession: H70846  
 A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-1538 <COL>  
 A: Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PID:e1299708; PID:g3261518  
 A: Experimental source: strain H37Rv  
 C: Genetics:  
 A: Gene: RV3345c

Query Match 61.0%; Score 36; DB 2; Length 1538;  
 Best Local Similarity 60.0%; Pred. No. 84;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RTRCRFKRRC 10  
 | | | | |

DB 1473 RRCRQORRC 1482

RESULT 4  
 R3EG14  
 ribosomal protein S14, chloroplast - Euglena gracilis chloroplast  
 C: Species: chloroplast Euglena gracilis  
 C: Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 30-Jan-1998  
 C: Accession: S04701; S34531; S34898  
 R: Nickoloff, J.A.; Christopher, D.A.; Drager, R.G.; Hallick, R.B.  
 Nucleic Acids Res. 17, 4882, 1989  
 A: Title: Nucleotide sequence of the Euglena gracilis chloroplast genes for isoleucine, P  
 A: Reference number: S04701; MUID: 89315233  
 A: Accession: S04701  
 A: Molecule type: DNA  
 A: Residues: 1-100 <NIC>  
 A: Cross-references: EMBL: X15240; NID: g11528; PID: g11529  
 A: Experimental source: strain 2  
 R: Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.; Monfort, A.; Orsat, B.; Spielman  
 submitted to the EMBL Data Library, January 1993  
 A: Description: The complete sequence of the Euglena gracilis chloroplast genome (tentati  
 A: Reference number: S34494  
 A: Accession: S34531  
 A: Molecule type: DNA  
 A: Residues: 1-100 <HALI>  
 A: Cross-references: EMBL: X70810; NID: g415327; PID: g415766  
 R: Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.R.; Monfort, A.; Orsat, B.; Spielman  
 Nucleic Acids Res. 21, 3537-3544, 1993  
 A: Title: Complete sequence of Euglena gracilis chloroplast DNA.  
 A: Reference number: S34862; MUID: 93347989  
 A: Accession: S34898  
 A: Status: nucleic acid sequence not shown; translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-100 <HAL2>  
 A: Cross-references: EMBL: X70810; NID: g415327; PID: g415766  
 A: Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
 C: Genetics:  
 A: Gene: rps14  
 A: Genome: chloroplast  
 A: Introns: 59/3  
 C: Superfamily: Escherichia coli ribosomal protein S14  
 C: Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 59.3%; Score 35; DB 1; Length 100;  
 Best Local Similarity 62.5%; Pred. No. 15;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRFRKRC 10  
 | | | |

DB 56 RCRLNRNC 63

RESULT 5

VHXRPV

minor outer capsid protein - porcine rotavirus C (strain OSU)  
 C: Species: porcine rotavirus C  
 C: Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 22-Jan-1999  
 C: Accession: A34009  
 R: Gonzalez, S.A.; Burrone, O.R.  
 Nucleic Acids Res. 17, 6402, 1989  
 A: Title: Porcine OSU rotavirus segment II sequence shows common features with the vir  
 A: Reference number: A34009; MUID: 89366681  
 A: Accession: A34009  
 A: Molecule type: genomic RNA  
 A: Residues: 1-197 <GON>  
 A: Cross-references: GB: X15519; NID: g61882; PID: g61883  
 C: Genetics:  
 A: Map position: segment 11  
 C: Superfamily: rotavirus minor outer capsid protein  
 C: Keywords: capsid protein; coat protein

Query Match 59.3%; Score 35; DB 1; Length 197;  
 Best Local Similarity 57.1%; Pred. No. 25;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRFKRRC 10

| | | | |

DB 173 CKYKRC 179

RESULT 6

D61014

defensin NP-1 - rat

N: Alternate names: cationic cysteine-rich peptide R-4

C: Species: Rattus norvegicus (Norway rat)

C: Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999

C: Accession: D61014; A60113

R: Belcourt, D.; Singh, A.; Bateman, A.; Lazure, C.; Solomon, S.; Bennett, H.P.J.

Regul. Pept. 40, 87-100, 1992

A: Title: Purification of cationic cysteine-rich peptides from rat bone marrow. Primary

A: Reference number: A61014; MUID: 93067247

A: Accession: D61014

A: Molecule type: protein

A: Residues: 1-32 <BEL>

R: Eisenhauer, P.B.; Harwig, S.L.; Szklarek, D.; Ganz, T.; Selsted, M.E.; Lehrer, R.I.

Infect. Immun. 57, 2021-2027, 1989

A: Title: Purification and antimicrobial properties of three defensins from rat neutro

A: Reference number: A60113; MUID: 89277517

A: Accession: A60113

A: Molecule type: protein

A: Residues: 1-32 <EIS>

C: Comment: Defensins are antimicrobial cationic peptides with activity against fungi

C: Superfamily: mammalian defensin

Query Match 59.3%; Score 35; DB 2; Length 32;  
 Best Local Similarity 66.7%; Pred. No. 6;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RTRCRFKRR 9

| | | | |

DB 7 RTRCGFRER 15

RESULT 7

JC4009

hepatocyte nuclear factor 4A - human

C: Species: Homo sapiens (man)

C: Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 31-Oct-1997

C: Accession: JC4009

R: Chartier, F.L.; Bossu, J.P.; Laudet, V.; Fruchart, J.C.; Laine, B.

**THIS PAGE BLANK (USPTO)**

-Wed Sep 8 09:42:38 1999

us-09-124-280a-17.ra1

Page 7

TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
US-08-158-189-31

Query Match 59.3%; Score 35; DB 1; Length 32;  
Best Local Similarity 66.7%; Pred. No. 5;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RTRCRFKRR 9  
| | | | | : |  
Db 7 RTRCGFRER 15

Search completed: September 7, 1999, 22:38:38  
Job time: 7922 sec



Query Match 62.7%; Score 37; DB 2; Length 11;  
Best Local Similarity 40.0%; Pred. No. 0.95;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RTCRFRKRC 10  
:1:1:1:1:  
DB 2 KTKCKFLKRC 11

## RESULT 13

US-08-456-112B-31  
; Sequence 31, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-31

Query Match 62.7%; Score 37; DB 2; Length 10;  
Best Local Similarity 40.0%; Pred. No. 0.88;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RTCRFRKRC 10  
:1:1:1:1:1:  
DB 1 KTKCKFLKRC 10

## RESULT 14

US-08-456-112B-35  
; Sequence 35, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York

COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-35

Query Match 62.7%; Score 37; DB 2; Length 11;  
Best Local Similarity 40.0%; Pred. No. 0.95;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RTCRFRKRC 10  
:1:1:1:1:1:  
DB 2 KTKCKFLKRC 11

## RESULT 15

US-08-158-189-31  
; Sequence 31, Application US/08158189  
; Patent No. 5641497  
; GENERAL INFORMATION:  
; APPLICANT: Bevins, Charles L.  
; TITLE OF INVENTION: Gastrointestinal Defensin Peptides,  
; TITLE OF INVENTION: CDNA Sequences, Methods for Production and Use Thereof  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5641497r1s  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/158,189  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/888,232  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Johnson, Philip S.  
; REGISTRATION NUMBER: 27,200  
; REFERENCE/DOCKET NUMBER: CH-0219  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100

ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/218,026  
APPLICATION NUMBER: US/08/218,026  
FILING DATE: 24-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalczyk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 600.286US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-9081  
TELEFAX: 612-332-5300  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
MOLECULE TYPE: peptide  
US-08-218-026-50

Query Match 62.7%; Score 37; DB 2; Length 11;  
Best Local Similarity 40.0%; Pred. No. 0.95;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 RTGCRFRRC 10  
:|:|:|:|:  
Db 2 KTKCKFLKCC 11

RESULT 11  
US-08-653-632-49  
Sequence 49, Application US/08653632  
Patent No. 5830860  
GENERAL INFORMATION:  
APPLICANT: GRAY, Beulah  
APPLICANT: HASEMAN, Judith R.  
APPLICANT: MAYO, Kevin  
TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 5830860west Center, 90 South seventh st  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/653,632  
FILING DATE: 24-MAY-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/218026  
FILING DATE: 24-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Kettelberger, Denise M  
REGISTRATION NUMBER: 33,924  
REFERENCE/DOCKET NUMBER: 600.286US11  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612/371-5268  
TELEFAX: 612/332-9081  
TELEX:  
INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-653-632-49

Query Match 62.7%; Score 37; DB 2; Length 10;  
Best Local Similarity 40.0%; Pred. No. 0.88;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 RTGCRFRRC 10  
:|:|:|:|:  
Db 1 KTKCKFLKCC 10

RESULT 12  
US-08-653-632-50  
Sequence 50, Application US/08653632  
Patent No. 5830860  
GENERAL INFORMATION:  
APPLICANT: GRAY, Beulah  
APPLICANT: HASEMAN, Judith R.  
APPLICANT: MAYO, Kevin  
TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 5830860west Center, 90 South seventh st  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/653,632  
FILING DATE: 24-MAY-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/218026  
FILING DATE: 24-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Kettelberger, Denise M  
REGISTRATION NUMBER: 33,924  
REFERENCE/DOCKET NUMBER: 600.286US11  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612/371-5268  
TELEFAX: 612/332-9081  
TELEX:  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-653-632-50

REFERENCE/DOCKET NUMBER: 576-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-2

Query Match 62.7%; Score 37; DB 1; Length 10;  
Best Local Similarity 40.0%; Pred. No. 0.88;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 RTCRFRRC 10  
:1:1:1:1:  
Db 1 KTKCKFLKRC 10

RESULT 8  
US-08-280-397-6  
Sequence 6, Application US/08280397  
Patent No. 5589459  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280.397  
FILING DATE: 07/26/94  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-6

Query Match 62.7%; Score 37; DB 1; Length 11;  
Best Local Similarity 40.0%; Pred. No. 0.95;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 RTCRFRRC 10  
:1:1:1:1:1:  
Db 2 KTKCKFLKRC 11

1

RESULT 9  
US-08-218-026-49  
Sequence 49, Application US/08218026  
Patent No. 5786324  
GENERAL INFORMATION:  
APPLICANT: Gray, Beulah  
APPLICANT: Haseman, Judith R.  
APPLICANT: Mayo, Kevin  
TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
TITLE OF INVENTION: Activity and Endotoxin Neutralizing Activity for Gram  
TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5786324west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/218,026  
FILING DATE: 24-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalchuk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 600.286US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-9081  
TELEFAX: 612-332-5300  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-218-026-49

Query Match 62.7%; Score 37; DB 2; Length 10;  
Best Local Similarity 40.0%; Pred. No. 0.88;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 RTCRFRRC 10  
:1:1:1:1:1:  
Db 1 KTKCKFLKRC 10

RESULT 10  
US-08-218-026-50  
Sequence 50, Application US/08218026  
Patent No. 5786324  
GENERAL INFORMATION:  
APPLICANT: Gray, Beulah  
APPLICANT: Haseman, Judith R.  
APPLICANT: Mayo, Kevin  
TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
TITLE OF INVENTION: Activity and Endotoxin Neutralizing Activity for Gram  
TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5786324west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA

OY 1 RTCRFRRC 10  
:1:1:1:1  
Db 2 KTKCKFLKRC 11

## RESULT 5

US-07-819-893-2  
; Sequence 2, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8989  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: Circular  
; US-07-819-893-2

Query Match 62.7%; Score 37; DB 1; Length 10;  
Best Local Similarity 40.0%; Pred. No. 0.88; 1; Indels 0; Gaps 0;  
Matches 4; Conservative 5; Mismatches 1;

OY 1 RTCRFRRC 10  
:1:1:1:1  
Db 1 KTKCKFLKRC 10

## RESULT 6

US-07-819-893-6  
; Sequence 6, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/819,893  
FILING DATE: 19920115  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8989  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: Circular  
US-07-819-893-6

Query Match 62.7%; Score 37; DB 1; Length 11;  
Best Local Similarity 40.0%; Pred. No. 0.95;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 RTCRFRRC 10  
:1:1:1:1  
Db 2 KTKCKFLKRC 11

## RESULT 7

US-08-280-397-2  
; Sequence 2, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hedman, Gibson & Costigan
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: LEADING EDGE 486
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,112B
FILING DATE: May 31, 1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Costigan, James V.
REGISTRATION NUMBER: 25,669
REFERENCE/DOCKET NUMBER: 576-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 302-8989
TELEFAX: (212) 302-8998
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: circular
US-08-456-112B-17

Query Match      100.0%; Score 59; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RTCRFRKRC 10
Db      1 RTCRFRKRC 10

RESULT      3
US-08-049-871-2
Sequence 2, Application US/08049871
Patent No. 5358933
GENERAL INFORMATION:
APPLICANT: Porro, Massimo
TITLE OF INVENTION: Synthetic Peptides for Detoxification
TITLE OF INVENTION: Of Bacterial Endotoxins and for the
TITLE OF INVENTION: Prevention and Treatment of Septic
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hedman, Gibson, Costigan & Hoare
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/049,871
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Costigan, James V.
REGISTRATION NUMBER: 25,669
REFERENCE/DOCKET NUMBER: 576-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 302-8989
TELEFAX: (212) 302-8998
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: circular
US-08-049-871-6

Query Match      62.7%; Score 37; DB 1; Length 11;
Best Local Similarity 40.0%; Pred. No. 0.95;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```

ATTORNEY/AGENT INFORMATION:
NAME: Costigan, James V.
REGISTRATION NUMBER: 25,669
REFERENCE/DOCKET NUMBER: 576-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 302-8989
TELEFAX: (212) 302-8998
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: circular
US-08-049-871-2

Query Match      62.7%; Score 37; DB 1; Length 10;
Best Local Similarity 40.0%; Pred. No. 0.88;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 RTCRFRKRC 10
Db      1 KTKCKFLKRC 10

RESULT      4
US-08-049-871-6
Sequence 6, Application US/08049871
Patent No. 5358933
GENERAL INFORMATION:
APPLICANT: Porro, Massimo
TITLE OF INVENTION: Synthetic Peptides for Detoxification
TITLE OF INVENTION: Of Bacterial Endotoxins and for the
TITLE OF INVENTION: Prevention and Treatment of Septic
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hedman, Gibson, Costigan & Hoare
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/049,871
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/658,744
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Costigan, James V.
REGISTRATION NUMBER: 25,669
REFERENCE/DOCKET NUMBER: 576-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 302-8989
TELEFAX: (212) 302-8998
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: circular
US-08-049-871-6

Query Match      62.7%; Score 37; DB 1; Length 11;
Best Local Similarity 40.0%; Pred. No. 0.95;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:37 ; Search time 80.79 Seconds  
(without alignments)

1.221 Million cell updates/sec

Title: US-09-124-280A-17

Perfect score: 59

Sequence: 1 RTCRFRKRC 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database:

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/PCTUS9\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	59	100.0	10	1	US-08-097-830E-17	Sequence 17, Appl
2	59	100.0	10	2	US-08-456-112B-17	Sequence 17, Appl
3	37	62.7	10	1	US-08-049-871-2	Sequence 2, Appl
4	37	62.7	10	1	US-08-049-871-6	Sequence 6, Appl
5	37	62.7	10	1	US-07-819-893-2	Sequence 2, Appl
6	37	62.7	10	1	US-07-819-893-6	Sequence 6, Appl
7	37	62.7	10	1	US-08-280-397-2	Sequence 6, Appl
8	37	62.7	10	1	US-08-280-397-6	Sequence 6, Appl
9	37	62.7	10	2	US-08-218-026-49	Sequence 49, Appl
10	37	62.7	10	2	US-08-218-026-50	Sequence 50, Appl
11	37	62.7	10	2	US-08-653-632-49	Sequence 49, Appl
12	37	62.7	10	2	US-08-653-632-50	Sequence 50, Appl
13	37	62.7	10	2	US-08-456-112B-31	Sequence 31, Appl
14	37	62.7	10	2	US-08-456-112B-35	Sequence 35, Appl
15	35	59.3	32	1	US-08-158-189-31	Sequence 31, Appl
16	35	59.3	32	2	US-08-491-204A-15	Sequence 15, Appl
17	33	55.9	10	1	US-08-472-219-1	Sequence 1, Appl
18	33	54.2	267	1	US-07-901-707-1	Sequence 7, Appl
19	32	54.2	48	1	US-07-965-674-7	Sequence 7, Appl
20	32	54.2	267	1	US-07-988-430-1	Sequence 1, Appl
21	32	54.2	267	1	US-08-218-303-16	Sequence 16, Appl
22	32	54.2	267	1	US-08-425-336-1	Sequence 1, Appl
23	32	54.2	290	1	US-08-378-761A-27	Sequence 27, Appl
24	32	54.2	540	1	US-08-378-761A-77	Sequence 77, Appl
25	32	54.2	290	1	US-08-485-286-27	Sequence 27, Appl
26	32	54.2	540	1	US-08-485-286-77	Sequence 77, Appl
27	32	54.2	10	1	US-08-097-830E-15	Sequence 15, Appl
28	32	54.2	267	1	US-08-488-113B-1	Sequence 1, Appl
29	32	54.2	48	1	US-08-462-661A-14	Sequence 14, Appl
30	32	54.2	48	1	US-08-462-661A-16	Sequence 16, Appl
31	32	54.2	267	2	US-08-477-484B-1	Sequence 1, Appl
32	32	54.2	10	2	US-08-456-112B-15	Sequence 15, Appl
33	32	54.2	267	2	US-08-646-360-1	Sequence 1, Appl
34	32	54.2	267	2	US-08-338-793D-61	Sequence 61, Appl
35	32	54.2	118	2	US-08-417-174-2	Sequence 2, Appl
36	32	54.2	63	2	US-08-369-829A-18	Sequence 18, Appl
37	32	54.2	118	2	US-08-231-565A-2	Sequence 2, Appl
38	32	54.2	268	2	US-08-356-786-8	Sequence 8, Appl
39	32	54.2	534	2	US-08-356-786-10	Sequence 10, Appl

40	32	54.2	267	3	PCT-US92-09487-1	Sequence 1, Appl
41	32	54.2	48	3	PCT-US93-09523-7	Sequence 7, Appl
42	30.5	51.7	802	1	US-07-862-021B-12	Sequence 12, Appl
43	30.5	51.7	802	2	US-08-313-288B-12	Sequence 12, Appl
44	30.5	51.7	802	3	PCT-US93-03164-12	Sequence 12, Appl
45	30	50.8	80	4	5223394-7	Patent No. 5223394

#### ALIGNMENTS

```
RESULT 1
US-08-097-830E-17
Sequence 17, Application US/08097830E
Patent No. 5652211
GENERAL INFORMATION:
APPLICANT: PORRO, Massimo
TITLE OF INVENTION: Peptides For Neutralizing The
NUMBER OF INVENTION: Toxicity of Lipid A
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: Hedman, Gibson & Costigan, P.C.
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,830E
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Costigan, James V.
REGISTRATION NUMBER: 25,669
TELEPHONE: (212) 302-8969
TELEFAX: (212) 302-8969
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: circular
FEATURE:
OTHER INFORMATION: sulfide bond between Cys and Cys
US-08-097-830E-17
Query Match 100.0%; Score 59; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 RTCRFRKRC 10
OY 1 RTCRFRKRC 10
|||||
DB 1 RTCRFRKRC 10
RESULT 2
US-08-456-112B-17
Sequence 17, Application US/08456112B
Patent No. 5834430
GENERAL INFORMATION:
APPLICANT: PORRO, Massimo
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
NUMBER OF SEQUENCES: 45
```

PD 19-MAR-1998.  
 PF 10-SEP-1997: U16037.  
 PR 30-OCT-1996: US-029679.  
 PR 10-SEP-1996: US-025719.  
 PR 02-OCT-1996: US-028056.  
 PA (ARCH-) ARCH DEV CORP.  
 PI Bell GI, Furuta H, Horikawa Y, Kaisaki PJ, Menzel S,  
 PI Oda N, Yamagata K;  
 DR WPI: 98-271667/24.  
 DR N-PSDB: V52731.  
 PT Isolated nucleic acid encoding hepatocyte nuclear factor 1-alpha and  
 PT 1-beta - useful for detecting susceptibility for non-insulin  
 PT dependent diabetes, especially maturity-onset diabetes of the young  
 PS Disclosure: Fig 28A-V: 363pp: English.  
 CC This is the amino acid sequence of human hepatocyte nuclear factor  
 CC 4 alpha (HNF-4 alpha) as deduced from a partial gene sequence (see  
 CC V52731). Alternative splicing results in a 516-amino acid  
 CC polypeptide (see W1587). The HNF-4 alpha sequence has also  
 CC been deduced from a cDNA clone (see W1574). HNF-4 alpha is a  
 CC transcription factor involved in regulating gene expression in  
 CC insulin-secreting beta cells. Mutations in HNF-4 alpha are  
 CC indicative of a propensity to diabetes mellitus. The invention  
 CC concerns the identification of genes responsible for non-insulin  
 CC dependent diabetes mellitus (NIDDM) for use in diagnostics and  
 CC therapeutics. It demonstrates that the MODY3 locus is the HNF-1  
 CC alpha gene (see V52625), the MODY4 locus is the HNF-1 beta gene  
 CC (see also V52730) and the MODY1 locus is the HNF-4 alpha gene (see  
 CC also V52687). Analysis of mutations in these HNF genes can be  
 CC diagnostic for diabetes. The invention also contemplates methods  
 CC of screening for modulators of HNF function utilising HNF nucleic  
 CC acids or polypeptides, the modulators being useful for treating  
 CC diabetes by modulating HNF function in an animal.  
 SO Sequence 567 AA;

Query Match 59.3%; Score 35; DB 1; Length 567;  
 Best Local Similarity 71.4%; Pred. NO. 1.5e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 CREKRC 10  
 |||||  
 Db 147 CREKRC 153

RESULT 15  
 W1587  
 ID W1587 standard: Protein: 516 AA.  
 AC W1587;  
 DT 21-DEC-1998 (first entry)  
 DE Human hepatocyte nuclear factor 4 alpha.  
 KM Hepatocyte nuclear factor 4 alpha: HNF-4 alpha: MODY1; human;  
 KM transcription factor; maturity onset diabetes of the young;  
 OS Homo sapiens.  
 PN MO9811254-A1.  
 PD 19-MAR-1998.  
 PE 10-SEP-1997: U16037.  
 PR 30-OCT-1996: US-029679.  
 PR 10-SEP-1996: US-025719.  
 PR 02-OCT-1996: US-028056.  
 PA (ARCH-) ARCH DEV CORP.  
 PI Bell GI, Furuta H, Horikawa Y, Kaisaki PJ, Menzel S,  
 PI Oda N, Yamagata K;  
 DR WPI: 98-271667/24.  
 DR N-PSDB: V52731.  
 PT Isolated nucleic acid encoding hepatocyte nuclear factor 1-alpha and  
 PT 1-beta - useful for detecting susceptibility for non-insulin  
 PT dependent diabetes, especially maturity-onset diabetes of the young  
 PS Disclosure: Fig 28A-V: 363pp: English.  
 CC This is the amino acid sequence of human hepatocyte nuclear factor  
 CC 4 alpha (HNF-4 alpha) as deduced from a partial gene sequence (see  
 CC V52731). Alternative splicing results in a 567-amino acid  
 CC polypeptide (see W1587). The HNF-4 alpha sequence has also

CC been deduced from a cDNA clone (see W1574). HNF-4 alpha is a  
 CC transcription factor involved in regulating gene expression in  
 CC insulin-secreting beta cells. Mutations in HNF-4 alpha are  
 CC indicative of a propensity to diabetes mellitus. The invention  
 CC concerns the identification of genes responsible for non-insulin  
 CC dependent diabetes mellitus (NIDDM) for use in diagnostics and  
 CC therapeutics. It demonstrates that the MODY3 locus is the HNF-1  
 CC alpha gene (see V52625), the MODY4 locus is the HNF-1 beta gene  
 CC (see also V52730) and the MODY1 locus is the HNF-4 alpha gene (see  
 CC also V52687). Analysis of mutations in these HNF genes can be  
 CC diagnostic for diabetes. The invention also contemplates methods  
 CC of screening for modulators of HNF function utilising HNF nucleic  
 CC acids or polypeptides, the modulators being useful for treating  
 CC diabetes by modulating HNF function in an animal.  
 SO Sequence 516 AA;

Query Match 59.3%; Score 35; DB 1; Length 516;  
 Best Local Similarity 71.4%; Pred. NO. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 CREKRC 10  
 |||||  
 Db 138 CREKRC 144

Search completed: September 7, 1999, 20:37:19  
 Job time: 18468 sec





Query Match 62.7%; Score 37; DB 1; Length 11;

Best Local Similarity 40.0%; Pred. No. 3.3;

Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RTRCRFKRC 10  
:|:|:|:|:|  
Db 2 KTKCKFLKKC 11

RESULT 9

W62434 ID W62434 standard; peptide: 10 AA.

AC W62434:

DT 01-OCT-1998 (first entry)

DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:49.

KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;

KM bactericidal permeability increasing factor; B/PI; neutralising;

KW human neutrophil granule bactericidal protein.

OS Homo sapiens.

OS Synthetic.

PN US5786324-A.

PD 28-JUL-1998.

PE 24-MAR-1994; 218026.

PR 24-MAR-1994; US-218026.

PA (MIND ) UNIV MINNESOTA.

PI Gray B. Haseman JR, Mayo K;

DR WPI: 98-436578/37.

PT Bactericidal and endotoxin-neutralising peptides - used in treating

PT e.g. Pseudomonas species infection and in protectively coating

PT prosthetic devices

PS Example 1: Column 18: 46pp; English.

CC The present invention describes bactericidal peptides (BP): (a) for

CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)

CC with both endotoxin neutralising activity and bactericidal activity for

CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising

CC activity and bactericidal activity, the peptide having at least 1 amino

CC acid sequence of a peptide with endotoxin neutralising activity combined

CC with at least 1 different amino acid sequence of a peptide that has

CC bactericidal activity; and (e) a prosthetic device which has a

CC sufficient amount of BP attached to the surface to inhibit bacterial

CC growth. The peptides of the invention are used in treating bacterial

CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7 to

CC 10-9 M, and Escherichia coli. The peptides are also used to treat

CC endotoxin shock. The present sequence represents a peptide derived

CC from human neutrophil granule bactericidal protein from an example of

CC the present invention.

CC Sequence 10 AA;

SO

Query Match 62.7%; Score 37; DB 1; Length 10;

Best Local Similarity 40.0%; Pred. No. 3.1;

Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RTRCRFKRC 10  
:|:|:|:|:|  
Db 1 KTKCKFLKKC 10

RESULT 10

W62435 ID W62435 standard; peptide: 11 AA.

AC W62435:

DT 01-OCT-1998 (first entry)

DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:50.

KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;

KM bactericidal permeability increasing factor; B/PI; neutralising;

KW human neutrophil granule bactericidal protein.

OS Homo sapiens.

OS Synthetic.

PN US5786324-A.

PD 28-JUL-1998.

PF 24-MAR-1994; 218026.

PR 24-MAR-1994; US-218026.

PA (MIND ) UNIV MINNESOTA.

PI Gray B. Haseman JR, Mayo K;

DR WPI: 98-436578/37.

PT Bactericidal and endotoxin-neutralising peptides - used in treating

PT e.g. Pseudomonas species infection and in protectively coating

PT prosthetic devices

PS Example 1: Column 18: 46pp; English.

CC The present invention describes bactericidal peptides (BP): (a) for

CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)

CC with both endotoxin neutralising activity and bactericidal activity for

CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising

CC activity and bactericidal activity, the peptide having at least 1 amino

CC acid sequence of a peptide with endotoxin neutralising activity combined

CC with at least 1 different amino acid sequence of a peptide that has

CC bactericidal activity; and (e) a prosthetic device which has a

CC sufficient amount of BP attached to the surface to inhibit bacterial

CC growth. The peptides of the invention are used in treating bacterial

CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7 to

CC 10-9 M, and Escherichia coli. The peptides are also used to treat

CC endotoxin shock. The present sequence represents a peptide derived

CC from human neutrophil granule bactericidal protein from an example of

CC the present invention.

SO Sequence 11 AA;

Query Match 62.7%; Score 37; DB 1; Length 11;

Best Local Similarity 40.0%; Pred. No. 3.3;

Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RTRCRFKRC 10  
:|:|:|:|:|  
Db 2 KTKCKFLKKC 11

RESULT 11

W77748 ID W77748 standard; Protein: 115 AA.

AC W77748:

DT 30-OCT-1998 (first entry)

DE Staphylococcus aureus protein of unknown function.

KW Staphylococcus aureus protein; immune response; gastrointestinal infection;

KW antibody production; T-cell immune response; gastroenteritis; infection;

KW respiratory infection; inhibitor; bacterial infection; cardiac infection;

KW central nervous system; kidney infection; urinary tract infection;

KW antimicrobial compound identification; broad spectrum antibiotic;

KW therapy.

OS Staphylococcus aureus.

OS Key Location/Qualifiers

FT Misc\_difference 1, 115

FT "residues designated X are unspecified, and

FT represented as Xaa in the specification"

EP-841394-A2.

PD 13-MAY-1998.

PE 24-SEP-1997; 307485.

PR 24-SEP-1996; US-027032.

PA (SMK ) SMITHKLINE BEECHAM CORP.

PA (SMK ) SMITHKLINE BEECHAM PLC.

PI Black MT, Burnham MKR, Hodgson JE, Knowles DJC,

PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M,

PI Ward JM;

DR WPI: 98-252940/23.

DR N-PSDB: V53536.

PT New nucleic acid sequences from Staphylococcus aureus WCHU29 -

PT useful in vaccines and for treatment of bacterial infections of e.g.

PT respiratory tract and central nervous system

PS Claim 11: Page 369; 390pp; English.

CC This sequence represents a Staphylococcus aureus protein of unknown

CC function, and is encoded by a DNA sequence of the invention.

CC The DNA sequences were isolated from Staphylococcus aureus WCHU29

CC (NCIMB 40771). Host cells containing the DNA sequences are used to

CC produce polypeptides or fragments. The proteins are used in the treatment

PR septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 4, Page 32; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns, etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 10 AA;

Query Match 62.7%; Score 37; DB 1; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 3.1;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RTKCRFKRRC 10  
 :1:1:1:1:1:  
 Db 1 KTKCKFLKRC 10

RESULT 6  
 R39292  
 ID R39292 standard; peptide; 11 AA.  
 AC R39292:  
 DE 22-DEC-1993 (first entry)  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 5..11  
 PN WO931415-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M:  
 DR WPI: 93-243143/30.  
 DE New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 8; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns, etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 62.7%; Score 37; DB 1; Length 11;  
 Best Local Similarity 40.0%; Pred. No. 3.3;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RTKCRFKRRC 10  
 :1:1:1:1:1:  
 Db 2 KTKCKFLKRC 11

RESULT 7  
 W21619  
 ID W21619 standard; peptide; 10 AA.  
 AC W21619:  
 DE 26-AUG-1997 (first entry)  
 KW Antibiotic potentiating peptide #31.  
 KW permeability; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M:  
 DR WPI: 97-034095/03.  
 DE Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 35; Page 27; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 62.7%; Score 37; DB 1; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 3.1;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RTKCRFKRRC 10  
 :1:1:1:1:1:  
 Db 1 KTKCKFLKRC 10

RESULT 8  
 W21623  
 ID W21623 standard; peptide; 11 AA.  
 AC W21623:  
 DE 26-AUG-1997 (first entry)  
 KW Antibiotic potentiating peptide #35.  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 5..11  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M:  
 DR WPI: 97-034095/03.  
 DE Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 39; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
CC Sequence 10 AA:

Query Match	100.0%	Score 59	DB 1	length 10
Best Local Similarity	100.0%	Pred. NC	0.0013	
Matches	10	Conservative	0	Indels 0
		Mismatches	0	Gaps 0

Qy	1	RTRCRFKRRC	10
Db	1	RTRCRFKRRC	10

RESULT 3  
R33526

ID R33526 standard: peptide; 10 AA.  
AC R33526;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
FH Key  
FH Location/Qualifiers  
FT disulfide bond 4..10  
FN ZAB200943-A.  
PD 25-NOV-1992.  
PE 10-FEB-1992: 000943.  
PR 11-FEB-1991: US-658744.  
PA (PORR/) PORRO M.  
PI Porro M;  
PI WPI: 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 4; Page 32; 39pp. English.

CC This is a specific example of a generic peptide of formula  
CC R-(Lys/Arg/Ser- Phe/Tyr/Trp- Leu/Ile/Val)<sub>n</sub>-R, where n = 1-100  
CC and each R is H, an amino acid residue or a fatty acid residue.  
CC The peptide is useful for treating or preventing septic shock,  
CC mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC µg/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotoxic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstab  
CC against proteases.  
SQ Sequence 10 AA:

Query Match	62.7%	Score 37	DB 1	Length 10
Best Local Similarity	40.0%	Pred. No. 3.1		
Matches 4	Conservative	5	Mismatches 1	Indels 0
			Gaps 0	

Qy	1	RTRCRFKRRRC	10
		: : :	:
Db	1	KTKCKFLKCC	10

## RESULT 4

R33531

AC	R33531:1993 (first entry)
DT	07-JUL-1993
DE	Peptide for treating septic shock.
KW	Toxic shock; blood endotoxin removal; serum; diagnostic reagent;
KW	Cytokine release control; treatment; pertussis; bacterial meningitis;
KW	HLI-related infections; polymyxin B.
OS	Synthetic.
Key	Location/Qualifiers
FT	disulfide_bond 5..11

PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.

DR WPI: 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 8: Page 32: 39pp: English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)-n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of Lipid A or  
CC lipopolysaccharide (LPS); and for treating peritonitis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of Lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 11 AA;

Query Match	62.7%	Score 37	DB 1	Length 11
Best Local Similarity	40.0%	Pred. No. 3.3		
Matches 4; Conservative		5; Mismatches	1; Indels	0; Gaps 0

```
QY      1  RTRCREFKRRC  10
          :|:|:|:|:|
Db      2  KTKCKFLKCC  11
```

RESULT	5
R39288	
ID	R39288 standard; peptide; 10 AA.
AC	R39288; 1993 (first entry)
DT	22-DEC-1993
DE	Endotoxin in Lipid A neutralising peptide.
KW	Bacterial endotoxin; non-toxic; treatment; toxic shock
KW	polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;
KW	removal; sera; vaccines; drug solns.; solutions; cytokine release;
KW	control; in vivo; in vitro; detoxification; detection;
KW	quantification.
OS	Synthetic.
EH	Key
ET	disulfide-bond 4. 10
PN	WO931415-A.
PD	22-JUL-1993.
PF	14-MAY-1992; E01060.
PR	16-JAN-1992; US-819893.
PA	(PORR/) PORRO M.
PI	PORRO M;
DR	WPt: 93-243143/30.
PT	New peptide(s) which neutralise lipid A of bacterial endotoxin -
FT	forming non-toxic, antigenic complex, used to treat or prevent

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37.18 : Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280a-17  
Perfect score: 59  
Sequence: 1 RTCRFRKRC 10

Scoring table: BIOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	59	100.0	10	1	R71788	Peptide neutralist
2	59	100.0	10	1	W21605	Antibiotic potent
3	37	62.7	10	1	R33526	Peptide for treati
4	37	62.7	11	1	R33531	Endotoxin lipid A
5	37	62.7	10	1	R33288	Endotoxin lipid A
6	37	62.7	11	1	R39292	Antibiotic potenti
7	37	62.7	10	1	W21619	Antibiotic potenti
8	37	62.7	11	1	W21623	Human neutrophil g
9	37	62.7	11	1	W62434	Human neutrophil g
10	37	62.7	11	1	W62435	Staphylococcus aur
11	37	62.7	11	1	W77748	Hepatocyte nuclear
12	35	59.3	455	1	R28757	Human native hepat
13	35	59.3	465	1	W71574	Human hepatocyte n
14	35	59.3	567	1	W71582	Human hepatocyte n
15	35	59.3	516	1	W71587	Human hepatocyte n
16	35	59.3	32	1	W66430	Cationic peptide R
17	35	59.3	32	1	W73207	Rat defensin-2 pep
18	34	57.6	51	1	W33694	Mouse protamine 1
19	34	57.6	301	1	W33695	Anti-human SC sling
20	33	55.9	201	1	W37085	Bovine bone calcif
21	33	55.9	201	1	R08130	Human bone calcif
22	33	55.9	201	1	R08129	Human bone calcif
23	33	55.9	201	1	R10841	Human bone calcif
24	33	55.9	265	1	R14423	Hepatitis GB virus
25	33	55.9	3164	1	R94345	Kappa-selective op
26	33	55.9	10	1	W40208	Kappa-selective op
27	33	55.9	10	1	W31922	Mycobacterium spec
28	33	55.9	134	1	Y04956	Ricin D Modified r
29	33	55.9	574	1	P90079	DNA sequence of ri
30	32	54.2	574	1	P94793	Ricin A encoded by
31	32	54.2	332	1	P95639	Ricin A agglutinin A
32	32	54.2	280	1	P95648	Biosynthetic multi
33	32	54.2	200	1	P80164	Ricin A gene produ
34	32	54.2	332	1	R06554	Sequence of Ricin
35	32	54.2	534	1	P70324	Sequence of Ricin
36	32	54.2	574	1	P70325	Sequence of Ricin
37	32	54.2	574	1	P70325	Sequence of Ricin
38	32	54.2	576	1	P70326	Sequence of Ricin
39	32	54.2	332	1	P70097	Ricin A. N-termina
40	32	54.2	48	1	R10226	Platelet aggregati
41	32	54.2	49	1	R12891	Eristostatib fibril
42	32	54.2	565	1	P60240	Preprotricin. New b
43	32	54.2	565	1	P60240	Preprotricin. New b

44 32 54.2 565 1 P50166  
45 32 54.2 174 1 Y04916

## ALIGNMENTS

```

RESULT 1
R71788
ID R71788 standard; peptide; 10 AA.
AC R71788:
DT 01-OCT-1995 (first entry)
DE Peptide neutralising toxicity of Lipid A.
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide_bond 4..10
PN WO9503327-A.
PD 02-FEB-1995; E02413.
PF 21-JUL-1994; US-097830.
PR 26-JUL-1993; US-097830.
PA (BIOS-); BIOSYNTH SRL.
PI Porto M;
DR WPI: 95-075190/10.
PT New peptide(s) for neutralising LPS endotoxin - comprising
repeating units of a basic aminoacid or basic and hydrophobic
PT amino acids
PS Claim 19: Page 21: 26pp; English.
CC New peptides are claimed which are linear or cyclic peptides which
CC include units of formula: (A)n, where A is the cationic amino acid Lys
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or
CC Trp, and p is 2 or greater.
CC The peptides bind to lipid A of endotoxins and do not exhibit haemolytic
CC activity. Hence they can be used therapeutically to treat septic shock
CC and also in vitro to detoxify vaccines, drug solutions, injectable
CC nutrient solutions, etc.
CC The present sequence is a specifically claimed example of the new
CC peptides.
SO Sequence 10 AA:

Query Match 100.0%; Score 59; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RTCRFRKRC 10
DB 1 RTCRFRKRC 10

RESULT 2
ID W21605 standard; peptide; 10 AA.
AC W21605:
DT 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #17.
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;
permeability; outer bacterial membrane.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide_bond 4..10
PN WO9638163-A1.
PD 05-DEC-1996; E02313.
PF 29-MAY-1996; US-456112.
PR 31-MAY-1995; US-456112.
PA (BIOS-); BIOSYNTH SRL.
PI Porto M; Varra M;
DR WPI: 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 21: Page 25; 37pp; English.

```

OY 1 KLKCKLKLC 10  
 ||| |||  
 Db 2 KLKSKLLSC 11

## RESULT 15

ID Q06679 PRELIMINARY; PRT; 776 AA.  
 AC Q06679;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
 DE SIMILAR TO G-PROTEIN BETA SUBUNITS.  
 GN D9798.12.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA JOHNSON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,  
 RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,  
 RA HALLSWORTH K., HAWKINS J., HILLER L., JIER M., JOHNSON D.,  
 RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,  
 RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,  
 RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,  
 RA WILSON R., WATERSTON R.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA DU Z.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA WATERSTON R.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA JIA Y., CHERRY J.M.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U32517; AAB64760.1; -;  
 SQ SEQUENCE 776 AA; 87801 MW; BB58514C CRC32;

## Query Match

Best Local Similarity 60.0%; Score 33; DB 3; Length 776;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 KLKCKLKLC 9  
 ||| |||  
 Db 415 KLKCKLTLK 423

Search completed: September 7, 1999, 20:34:48  
 Job time: 19744 sec

OC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 96337999.  
 RA BUT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,  
 RA SCOTT J.L., GEORGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKIY M.,  
 RA KLEIN H.-P., FRASER C.M., SMITH H.O., MOESE C.R., VENTER J.C.,  
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RA jannaschii.";  
 RT Science 273:1058-1073(1996).  
 RL -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO 4FE-4S  
 CC FERREDOXINS.  
 DR EMBL: U67537; AAB98939.1; -;  
 DR PROSITE: PS00198; 4FEAS\_FERREDOXIN; 5.  
 DR PRAM: PF00037; fer4; 4.  
 DR Hypothetical protein; 4fe-4s.  
 KW Hypothetical protein; 4fe-4s. (BY SIMILARITY).  
 FT METAL 47 47 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 FT METAL 50 50 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 FT METAL 53 53 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 FT METAL 57 57 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
 FT METAL 99 99 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
 FT METAL 102 102 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
 FT METAL 105 105 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
 FT METAL 109 109 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 SQ SEQUENCE 209 AA; 23567 MW; 3FF664B3 CRC32.

Query Match 60.0%; Score 33; DB 1; Length 209;  
 Best Local Similarity 75.0%; Pred. No. 57;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 KCKLKLKC 10  
 Db 153 KCKLKLSC 160

RESULT 13  
 059556 PRELIMINARY; PRT: 170 AA.  
 AC 059556;  
 DT 01-AUG-1998 (TREMBLER, 07, Created)  
 DT 01-AUG-1998 (TREMBLER, 07, last sequence update)  
 DT 01-JAN-1999 (TREMBLER, 09, last annotation update)  
 DE 170AA LONG HYPOTHEICAL PROTEIN.  
 GN PH1864.  
 OS Pyrococcus horikoshii.  
 OC Archaea: Euryarchaeota; Thermococcales; Pyrococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OT3;  
 RX MEDLINE: 98344137.  
 RA KAMARABAYASI Y., SAMADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
 RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOGYAMA A., NAGAI Y.,  
 RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAKITA M., OHFURU Y.,  
 RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI Y., KUSHIDA N., OGUCHI A.,  
 RA AOKI K., NAKAMURA Y., ROSS T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
 RA KIKUCHI H.,  
 RA "Complete sequence and gene organization of the genome of a  
 RT hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL: AP000007; BAA30986.1; -;  
 SQ SEQUENCE 170 AA; 19813 MW; F9DEC0A1 CRC32;

Query Match 60.0%; Score 33; DB 1; Length 170;  
 Best Local Similarity 60.0%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 KCKLKLKC 10  
 Db 16 KCKLKLSC 25

RESULT 14  
 P71001 PRELIMINARY; PRT: 39 AA.  
 AC P71001;  
 DT 01-FEB-1997 (TREMBLER, 02, Created)  
 DT 01-FEB-1997 (TREMBLER, 02, last sequence update)  
 DT 01-NOV-1998 (TREMBLER, 08, last annotation update)  
 DE HYPOTHEICAL 4.2 KD PROTEIN.  
 GN PHF.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 CC Bacillus/staphylococcus group; Bacillus.  
 DR [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE: 98044033.  
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AZEVEDO V., BERBERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA BORRIS R., BOUSCHER L., BRANS A., BRAUN M., BRIGNELLI S.C., BRON S.,  
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPRANO V., CARTER N.M.,  
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRICH S.D., EMERSON P.T.,  
 RA ENTIAN K.D., ERRINGTON J., FABREY C., FERRARI E., FOULDER D.,  
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,  
 RA GHIM S.Y., GLASER P., GOFREAU A., GOLIGHTLY E.J., GRANDI A.,  
 RA GUISSEPT G., GUI B.J., HAGA K., HAIRCH J., HARWOOD C.R., JONES L.,  
 RA HILBERT H., HOLSAPPEL S., HOSONO S., KLERR-BLANCARD M., KLEIN C.,  
 RA JORIS B., KARAWATA D., KASAHARA Y., KLERR-BLANCARD M., KLEIN C.,  
 RA KOBAYASHI Y., KOETTER P., KONINGSSTEIN G., KROCH S., KUMANO M.,  
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
 RA MEDINA N., MELLANO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
 RA NOONE D., O'REILLY M., OGAMA K., OGIMARA A., OUDESA B., PARK S.H.,  
 RA PARRO V., POHL T.M., PORTERLE D., PORROLIK S., PRESCOTT A.M.,  
 RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
 RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADIE Y.,  
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOPFONE F.,  
 RA SEKICUCHI J., SEKOWSKA A., SEROR S.J., SEROR P., SHIN B.S., SOLDO B.,  
 RA SOROKIN A., TACCONE E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
 RA TAKEUCHI M., TAMAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
 RA TOSATO V., UCHITAYA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
 RA VIARI A., WAMBUTT R., WEDLER E., WEDLER H., WETZENEGER T., YATA K.,  
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO H., DANCIN A.,  
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCIN A.,  
 RT "The complete genome sequence of the gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL Nature 390:249-256(1997).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCIN A.,  
 RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z80360; CAB02499.1; -;  
 DR EMBL: Z99123; CAB15774.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 39 AA; 4199 MW; 896597AC CRC32;

Query Match 60.0%; Score 33; DB 2; Length 39;  
 Best Local Similarity 70.0%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

RESULT 9
ID 022328 PRELIMINARY: PRT: 560 AA.
AC 022328;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE COSMID T07H6.
GN T07H6.5.
OS Ctenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BOFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RT Nature 368:32-38(1994).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA GEISEL C.;
RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RL [3]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U53344; AAA96225.1; -.
DR PFM; PF00084; sush1.7.
SQ SEQUENCE 560 AA; 61619 MW; 3A10AF63 CRC32;

Query Match 61.8%; Score 34; DB 5; Length 560;
Best Local Similarity 71.4%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 CKLKLKC 10
Db 71 CKLKLKC 77

RESULT 10
ID 081330 PRELIMINARY: PRT: 991 AA.
AC 081330;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE F3D13.1 PROTEIN.
GN F3D13.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidops.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA WASHU;

```

```

RT "The A. thaliana Genome Sequencing Project."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA CORDES M., WOLLAM C., PAPE K.;
RT "The sequence of A. thaliana F3D13."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA WATERSTON R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF069300; AAC19289.1; -.
SQ SEQUENCE 991 AA; 11542 MW; 7E08A5D9 CRC32;

Query Match 61.8%; Score 34; DB 10; Length 991;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KLCCKLKLK 9
Db 281 KLCCKLKLK 289

RESULT 11
ID 082745 PRELIMINARY: PRT: 1773 AA.
AC 082745;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HYPOTHETICAL 199.0 KD PROTEIN.
GN F7H19.150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidops.
RN [1]
RP SEQUENCE FROM N.A.
RA PETERS S.A., VAN STRAVEREN M., DIRKSE W., STIEKEMA W., BANCROFT I.,
RA MIMES H.W., MAYER K.F.X., SCHUELLER C., BEVAN M.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU ARABIDOPSIS SEQUENCING PROJECT;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031018; CAA19812.1; -.
KV Hypothetical protein.
SQ SEQUENCE 1773 AA; 199005 MW; 53E18746 CRC32;

Query Match 61.8%; Score 34; DB 10; Length 1773;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 LKCKLKLK 9
Db 974 VKCKLKLK 981

RESULT 12
ID 058344 PRELIMINARY: PRT: 209 AA.
AC 058344;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HYPOTHETICAL PROTEIN M00934.
GN M00934.
OS Methanococcus jannaschii.

```

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RN SEQUENCE FROM N.A.  
 RP MPTS;  
 RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: U53880; AAB67591.1; -;  
 DR EMBL: 273259; CAA97647.1; -;  
 DR SGD: L0000557; GAA1.  
 SQ SEQUENCE 2958 AA; 338254 MW; 60D78E22 CRC32;

Query Match 63.6%; Score 35; DB 3; Length 2958;  
 Best Local Similarity 50.0%; Pred. No. 2e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KLCCKLKLC 10  
 |||:|:|:  
 Db 126 KLCCKISVEC 135

RESULT 6  
 ID 021440 PRELIMINARY; PRT: 1992 AA.  
 AC 021440;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE MYO-3 PROTEIN.  
 GN MYO-3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,  
 RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HARKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHONKKEEN R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 278199; CAB01576.1; -;  
 DR PFAM: PF00612; IO: 1.  
 DR PFAM: PF00063; myosin\_head; 2.  
 SQ SEQUENCE 1992 AA; 228362 MW; D0011BCA CRC32;

Query Match 63.6%; Score 35; DB 5; Length 1992;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KLCCKLKLC 8  
 |||:|:|:  
 Db 1906 KLCCKLKLC 1913

RESULT 7  
 ID 046008 PRELIMINARY; PRT: 2356 AA.  
 AC 046008;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE ZK228.2 PROTEIN.  
 GN ZK228.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA BASHAM V.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,  
 RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HARKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHONKKEEN R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 282086; CAB04994.1; -;  
 DR PFAM: PF00078; rvl; 1.  
 SQ SEQUENCE 2356 AA; 264212 MW; 0F6DDEBF CRC32;

Query Match 63.6%; Score 35; DB 5; Length 2356;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LKCKLKLC 10  
 :|:|:|:  
 Db 789 VKCKCKLVC 797

RESULT 8  
 ID 013999 PRELIMINARY; PRT: 187 AA.  
 AC 013999;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last annotation update)  
 DE HYPOTHETICAL 22.4 KD PROTEIN C27E2.04C IN CHROMOSOME I.  
 GN SPAC27E2.04C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archaescomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WOOD V.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: 298978; CAB1678.1; -;  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 94 114 POTENTIAL.  
 FT TRANSMEM 163 183 POTENTIAL.  
 FT DOMAIN 97 100 POLY-ILEU.  
 SQ SEQUENCE 187 AA; 22402 MW; 4E016BDC CRC32;

Query Match 61.8%; Score 34; DB 3; Length 187;  
 Best Local Similarity 40.0%; Pred. No. 35;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 KLCCKLKLC 10  
 :|:|:|:  
 Db 60 EYCKLKLC 69



DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)  
 DE U03088.  
 OS Mycobacterium leprae.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ROBINSON K.;  
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA SMITH D.R.;  
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA ROBINSON K.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U00022; AAA17336.1; -.  
 DR PFAM; PF00408; PGM\_PMW; 1.  
 DR PROSITE; PS00710; PGM\_PMW; 1.  
 SQ SEQUENCE 534 AA; 56627 MW; 668B1C41 CRC32;

Query Match 69.1%; Score 38; DB 2; Length 534;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KKKCKLKLC 10  
 |||||  
 Db 498 KKKCYLCVRC 507

RESULT 3  
 ID 097068 PRELIMINARY; PRT; 2065 AA.  
 AC 097068;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)  
 DE TRANSCRIPTION FACTOR.  
 GN TAF250.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
 OC Drosophilidae; Drosophila.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA CELINKER S.E., PEIFFER B., KNAFELS J., MARTIN C.H., MAYEDA C.A.,  
 RA PALAZZOLO M.J.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; A0001574; AAD1815.1; -.  
 DR PROSITE; PS00633; BROMODOMAIN\_1; 2.  
 SQ SEQUENCE 2065 AA; 232220 MW; EB8344A1 CRC32;

Query Match 67.3%; Score 37; DB 5; Length 2065;  
 Best Local Similarity 80.0%; Pred. No. 68;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 KKKCKLKLC 10  
 |||||  
 Db 1339 KKKPDLKLC 1348

RESULT 4  
 ID 098854 PRELIMINARY; PRT; 361 AA.  
 AC 098854;  
 DT 01-FEB-1997 (TREMblrel. 02, Created)  
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)

DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)  
 DE TRANSFORMING GROWTH FACTOR-BETA 2 (FRAGMENT).  
 OS Cyprinus carpio (Common carp).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;  
 OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Cyprinidae; Cyprinidae; Cyprinus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE; 97354301.  
 RA SUMATHY K., DESAI K.V., KONDALIAH P.;  
 RT "Isolation of transforming growth factor-beta2 cDNA from a fish,  
 RT Cyprinus carpio by RT-PCR."  
 CC Gene 191:103-107(1997).  
 CC -I- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.  
 DR EMBL; U66874; AAB62983.1; -.  
 DR PFAM; PF00019; TGF-beta; 1.  
 DR PFAM; PF00688; TGF-beta; 1.  
 DR PROSITE; PS00250; TGF-BETA; 1.  
 KW Glycoprotein.  
 FT NON\_TER 1 1  
 FT NON\_TER 361 361  
 FT SEQUENCE 361 AA; 41931 MW; 908A56D8 CRC32;

Query Match 67.3%; Score 37; DB 13; Length 361;  
 Best Local Similarity 60.0%; Pred. No. 18;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KKKCKLKLC 10  
 ::|||  
 Db 1 QILCKLKISC 10

RESULT 5  
 ID 012150 PRELIMINARY; PRT; 2958 AA.  
 AC 012150;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMblrel. 01, Last annotation update)  
 DE CHROMOSOME XII COSMID 9449.  
 GN GAA1 OR L9449.14.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA GEISEL C.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,  
 RA FAVELLO A., FULTON L., GATUNG S., GRECO T., KRISTEN J., KUCABA T.,  
 RA HALSWORTH K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.,  
 RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,  
 RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,  
 RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,  
 RA WILSON R., WATERSTON R.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA WATERSTON R.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 2353-2958 FROM N.A.  
 RA POHL T.M.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA BENES V., RECHMANN S., NENTWICH U., SCHWAGER C., ANSORGE W., VOSS H.;

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:47 : Search time 148.39 seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-16

Perfect score: 55

Sequence: 1 KLKCKLKLC 10

Scoring table: BLOSUM62

Searched: 201082 segs, 61543640 residues

Database :

SPTREMBL\_10:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertibrate:\*  
14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	39	70.9	534	2	053360	053360 mycobacteri
2	38	69.1	534	2	049913	049913 mycobacteri
3	37	67.3	2065	5	097068	097068 drosophila
4	37	67.3	361	13	098854	098854 cyprinus ca
5	35	63.6	2958	3	012150	012150 saccharomyc
6	35	63.6	1992	5	021440	021440 caenorhabdi
7	35	63.6	2356	5	046008	046008 caenorhabdi
8	34	61.8	187	3	013999	013999 schizosacch
9	34	61.8	360	5	022328	022328 caenorhabdi
10	34	61.8	991	10	081330	081330 arabidopsis
11	34	61.8	1773	10	082745	082745 arabidopsis
12	33	60.0	170	1	058344	058344 methanococ
13	33	60.0	170	1	059556	059556 pyrococcus
14	33	60.0	39	2	P71001	P71001 bacillus su
15	33	60.0	776	3	006679	006679 saccharomyc
16	33	60.0	1101	3	094172	094172 pneumocysti
17	33	60.0	26926	4	010466	010466 homo sapien
18	33	60.0	336	5	022444	022444 caenorhabdi
19	33	60.0	224	5	P91567	P91567 caenorhabdi
20	33	60.0	1043	5	001757	001757 caenorhabdi
21	33	60.0	2539	5	096157	096157 plasmodium
22	33	60.0	3085	5	097283	097283 plasmodium
23	33	60.0	505	8	020277	020277 plumeria ru
24	33	60.0	948	10	081211	081211 arabidopsis
25	33	60.0	1129	11	062924	062924 rattus norv
26	33	60.0	661	11	061889	061889 mus musculu
27	33	60.0	5376	11	088799	088799 mus musculu
28	33	60.0	127	12	088623	088623 beet necrot
29	33	60.0	182	12	084455	084455 paramediciu

## ALIGNMENTS

30	33	60.0	129	12	065677	065677 beet necrot
31	32	58.2	136	1	028432	028432 archaeoglob
32	32	58.2	448	1	027309	027309 methanobact
33	32	58.2	205	2	092013	092013 rickettsia
34	32	58.2	432	3	060142	060142 schizosacch
35	32	58.2	764	3	060201	060201 aspergillus
36	32	58.2	1299	4	060343	060343 homo sapien
37	32	58.2	441	4	000149	000149 homo sapien
38	32	58.2	517	4	075530	075530 homo sapien
39	32	58.2	777	4	095513	095513 schistosoma
40	32	58.2	119	5	017455	017455 schistosoma
41	32	58.2	610	5	027324	027324 drosophila
42	32	58.2	271	5	077208	077208 lytechinus
43	32	58.2	930	5	096165	096165 plasmodium
44	32	58.2	420	11	P70619	P70619 rattus norv
45	32	58.2				

RESULT 1  
053360 PRELIMINARY; PRT; 534 AA.  
AC 053360;  
ID 053360;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE PUTATIVE PHOSPHOMANNOSE-6-PHOSPHATASE.  
GN MT016.07.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37RV;  
RA OLIVER K., HARRIS D.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37RV;  
RA COLE S.T., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37RV;  
RA MEDLINE: 96181548.  
RA PHILLIPS W.J., POULTER S., EIGMEIER K., PASCOPELLA L.,  
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,  
RA COLE S.T.;  
RT "An integrated map of the genome of the tubercle bacillus,  
RT Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium  
RT leprae.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).  
DR EMBL: AL021841; CAAL1080.1; -  
DR PFM: PFM0408; PGM\_PMM; 1.  
DR PROSITE: PS00710; PGM\_PMM; 1.  
SQ SEQUENCE 534 AA; 56195 MW; 4FEDE95B4 CRC32;

Query Match 70.9%; Score 39; DB 2; Length 534;  
Best local Similarity 60.0%; Pred. No. 11;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLKCKLKLC 10  
DB 497 KLKCKLKLC 506

RESULT 2  
ID 049913 PRELIMINARY; PRT; 534 AA.  
AC 049913;

CYAB\_HUMAN STANDARD: PRT; 1251 AA.  
 ID CYAB\_HUMAN  
 AC P40145;  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DE ADENYLATE CYCLASE, TYPE VIII (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE)  
 DE (CA(2+)/CALMODULIN ACTIVATED ADENYLYL CYCLASE).  
 GN ADCY8.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATHARTINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN STEM;  
 RX MEDLINE: 94357261.  
 RA DEBER N., MARIN O., STENGEL D., DANISOVA A., IOURGENKO V.,  
 RA MATSUOKA I., CAPUT D., HANOUNE J.;  
 RT "Molecular cloning of the human type VIII adenylyl cyclase.";   
 RL FEBS LETT. 351:109-113(1994).  
 RN [2]  
 RP SEQUENCE OF 577-1249 FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE: 91354291.  
 RA PARMA J., STENGEL D., GANNAGE M.H., POYARD M., BAROUKI R., HANOUNE J.;  
 RT "Sequence of a human brain adenylyl cyclase partial cDNA: evidence  
 for a consensus cyclase specific domain.";  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 179:455-462(1991)  
 CC -1- FUNCTION: THIS A MEMBRANE-BOUND, CA(2+)-INHIBITABLE ADENYLYL  
 CYCLASE. MAY BE INVOLVED IN LEARNING, IN MEMORY AND IN DRUG  
 DEPENDENCE.  
 CC -1- CATALYTIC ACTIVITY: ATP -> 3',5'-CYCLIC AMP + PYROPHOSPHATE.  
 CC -1- ENZYME REGULATION: ACTIVATED BY CA(2+)/CALMODULIN.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: THE TWO CYTOPLASMIC DOMAINS ARE SIMILAR TO ONE  
 ANOTHER AND TO THE CATALYTIC DOMAINS OF GUANYLATE CYCLASES.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
 CC  
 CC EMBL: Z53309; G516263; -;  
 DR EMBL: M83533; G178542; -;  
 DR MIM: 103070; -;  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES; 2.  
 DR PFAM: PF00211; guanylate\_cyc; 2.  
 DR HSSP: P19754; 1AMK.  
 KW LYSASE; CAMP SYNTHESIS; TRANSMEMBRANE; GLYCOPROTEIN; DUPLICATION.  
 FT DOMAIN 1 182 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 183 203 POTENTIAL.  
 FT TRANSSEM 212 232 POTENTIAL.  
 FT TRANSSEM 247 267 POTENTIAL.  
 FT TRANSSEM 274 294 POTENTIAL.  
 FT TRANSSEM 296 316 POTENTIAL.  
 FT TRANSSEM 321 341 POTENTIAL.  
 FT TRANSSEM 343 363 POTENTIAL.  
 FT DOMAIN 363 715 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 716 736 POTENTIAL.  
 FT TRANSSEM 738 758 POTENTIAL.  
 FT TRANSSEM 787 807 POTENTIAL.  
 FT TRANSSEM 831 851 POTENTIAL.  
 FT TRANSSEM 861 881 POTENTIAL.  
 FT TRANSSEM 894 914 POTENTIAL.  
 FT DOMAIN 915 1251 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 817 817 POTENTIAL.  
 FT CARBOHYD 821 821 POTENTIAL.  
 FT CARBOHYD 888 888 POTENTIAL.  
 SQ SEQUENCE 1251 AA; 140121 MW; 1DAD8693 CRC32;

Query Match 59.6%; Score 31; DB 1; Length 1251;  
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KSLSLRLTY 10  
 11: 11 11:  
 Db 811 KSLPLKRLTF 820  
 RESULT 15  
 CYAB\_MOUSE STANDARD: PRT; 1249 AA.  
 ID CYAB\_MOUSE  
 AC P97490;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DE ADENYLATE CYCLASE, TYPE VIII (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE)  
 DE (CA(2+)/CALMODULIN ACTIVATED ADENYLYL CYCLASE).  
 GN ADCY8.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-BRAIN;  
 RA PREMONT R.T.;  
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBI DATA BANKS.  
 CC -1- FUNCTION: THIS A MEMBRANE-BOUND, CA(2+)-INHIBITABLE ADENYLYL  
 CYCLASE. MAY BE INVOLVED IN LEARNING, IN MEMORY AND IN DRUG  
 DEPENDENCE.  
 CC -1- CATALYTIC ACTIVITY: ATP -> 3',5'-CYCLIC AMP + PYROPHOSPHATE.  
 CC -1- ENZYME REGULATION: ACTIVATED BY CA(2+)/CALMODULIN.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: THE TWO CYTOPLASMIC DOMAINS ARE SIMILAR TO ONE  
 ANOTHER AND TO THE CATALYTIC DOMAINS OF GUANYLATE CYCLASES.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
 CC  
 CC EMBL: U85021; G1814375; -;  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES; 2.  
 DR PFAM: PF00211; guanylate\_cyc; 2.  
 DR HSSP: P19754; 1AMK.  
 KW LYSASE; CAMP SYNTHESIS; TRANSMEMBRANE; GLYCOPROTEIN; DUPLICATION.  
 FT DOMAIN 1 180 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 181 201 POTENTIAL.  
 FT TRANSSEM 210 230 POTENTIAL.  
 FT TRANSSEM 245 265 POTENTIAL.  
 FT TRANSSEM 272 292 POTENTIAL.  
 FT TRANSSEM 294 314 POTENTIAL.  
 FT TRANSSEM 319 339 POTENTIAL.  
 FT DOMAIN 340 713 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 714 734 POTENTIAL.  
 FT TRANSSEM 736 756 POTENTIAL.  
 FT TRANSSEM 785 805 POTENTIAL.  
 FT TRANSSEM 829 849 POTENTIAL.  
 FT TRANSSEM 859 879 POTENTIAL.  
 FT TRANSSEM 892 912 POTENTIAL.  
 FT DOMAIN 913 1249 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 815 815 POTENTIAL.  
 FT CARBOHYD 819 819 POTENTIAL.  
 FT CARBOHYD 886 886 POTENTIAL.  
 SQ SEQUENCE 1249 AA; 140154 MW; DA14EE4E CRC32;

Query Match 59.6%; Score 31; DB 1; Length 1249;  
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RT the highly heterogeneous nature of the viral population.";  
 VIROLOG 167:33-45(1992).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M81059; G33576; -  
 CC DR PIR: A42189; YGVNDB.  
 CC DR PF00974; Rhabd.glycop. 1.  
 CC TRANSMEMBRANE; ENVELOPE PROTEIN; GLYCOPROTEIN; SIGNAL.  
 CC SIGNAL 1 19  
 CC FT CHAIN 20 524 SPIKE GLYCOPROTEIN.  
 CC FT TRANSMEM 460 476 POTENTIAL.  
 CC FT CARBOHYD 56 56 POTENTIAL.  
 CC FT CARBOHYD 177 177 POTENTIAL.  
 CC FT CARBOHYD 338 338 POTENTIAL.  
 CC SEQUENCE 524 AA; 58722 MW; 8AC2FCA2 CRC32;  
 SO  
 Query Match 61.5%; Score 32; DB 1; Length 524;  
 Best Local Similarity 60.0%; Pred. No. 31;  
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 KSLSKRLTY 10  
 Db 313 KSVSLRLSH 322  
 RESULT 12  
 YCBM\_BACSU STANDARD; PRT: 158 AA.  
 ID YCBM\_BACSU  
 AC P42245;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE HYPOTHEICAL 18.0 KD PROTEIN IN PCP 5' REGION (ORF12).  
 GN YCBM.  
 OS BACILLUS SUBTILIS  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 OC BACILLUS.  
 CC [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN:168;  
 CC RA OGAWA K.-I., AKAGAWA E., NAKAMURA K., YAMANE K.;  
 RT "Determination of a 21548 bp nucleotide sequence around the 24  
 RT degrees region of the Bacillus subtilis chromosome."  
 RL MICROBIOLOGY 141:269-275(1995).  
 CC -1- SIMILARITY: TO THE C-TERMINAL OF PROKARYOTIC SENSORY TRANSDUCTION  
 CC HISTIDINE KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D30808; G71007; -  
 CC DR EMBL: Z99105; E1182208; -  
 CC SUBTILIST; Bg11168; YCBM;  
 CC PFAM: PF00512; signal: 1.  
 CC HYPOTHEICAL PROTEIN.  
 CC SEQUENCE 158 AA; 17994 MW; D6DB3D78 CRC32;  
 SO  
 Query Match 61.5%; Score 32; DB 1; Length 158;

Best Local Similarity 55.6%; Pred. No. 8.1;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 SLSKRLTY 10  
 Db 150 TITLRLTY 158  
 RESULT 13  
 ARP\_ARATH STANDARD; PRT: 527 AA.  
 ID ARP\_ARATH  
 AC P45951;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE APURINIC ENDONUCLEASE-REDOX PROTEIN (DNA-(APURINIC OR APYRIMIDINIC  
 DE SITE) LYASE) (EC 4.2.99.18) (FRAGMENT).  
 GN ARP OR REF.  
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
 OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 OC CAPRARIACE; BRASSICACEAE; ARABIDOPSIS.  
 CC [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=CV. COLUMBIA; TISSUE=CALLUS;  
 CC MEDLINE: 94211851.  
 CC RA BABYCHUK E., KUSHNIR S., VAN MONTAGU M., INZE D.;  
 RT "The Arabidopsis thaliana apurinic endonuclease Arp reduces human  
 RT transcription factors Fos and Jun."  
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:3299-3303(1994).  
 CC -1- FUNCTION: REPAIRS OXIDATIVE DNA DAMAGES. SEEMS ALSO TO ACT AS A  
 CC REDOX FACTOR. IS MULTIFUNCTIONAL AND MAY BE INVOLVED BOTH IN DNA  
 CC REPAIR AND IN THE REGULATION OF TRANSCRIPTION.  
 CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE NEAR APURINIC OR  
 CC APYRIMIDINIC SITES TO PRODUCTS WITH 5'-PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE STILONES, FLOWERS, AND STEMS.  
 CC -1- A HIGH LEVEL EXPRESSION IS SEEN IN THE LEAVES.  
 CC -1- SIMILARITY: BELONGS TO THE AP/EXO FAMILY OF DNA REPAIR ENZYMS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X76912; G472869; -  
 CC DR PROSITE; PS00726; AP\_NUCLEASE\_F1\_1; 1.  
 CC DR PROSITE; PS00727; AP\_NUCLEASE\_F1\_2; 1.  
 CC DR PROSITE; PS00728; AP\_NUCLEASE\_F1\_3; 1.  
 CC PFAM: PF01260; AP\_endonuclease1; 1.  
 CC KW DNA REPAIR; ENDONUCLEASE; HYDROLASE; NUCLEASE; LYASE; NUCLEAR PROTEIN.  
 CC FT NON TER 1 1  
 CC FT DOMAIN <1 269 HIGHLY CHARGED; INCREASES THE AFFINITY OF  
 CC ARP FOR DNA.  
 CC FT DOMAIN 270 527 AP ENDONUCLEASE.  
 CC FT METAL 304 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT ACT SITE 518 GENERAL BASE (BY SIMILARITY).  
 CC SEQUENCE 527 AA; 59243 MW; 52EDA79F CRC32;  
 SO  
 Query Match 59.6%; Score 31; DB 1; Length 527;  
 Best Local Similarity 85.7%; Pred. No. 50;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 5 LKRLTYR 11  
 Db 388 LKRLSTR 394  
 RESULT 14

FT DISULFID 2739 2759 BY SIMILARITY.  
 FT DISULFID 2761 2779 BY SIMILARITY.  
 FT DISULFID 2792 2805 BY SIMILARITY.  
 FT DISULFID 2800 2818 BY SIMILARITY.  
 FT DISULFID 2812 2829 BY SIMILARITY.  
 FT DISULFID 2834 2846 BY SIMILARITY.  
 FT DISULFID 2841 2859 BY SIMILARITY.  
 FT DISULFID 2853 2868 BY SIMILARITY.  
 FT DISULFID 2874 2886 BY SIMILARITY.  
 FT DISULFID 2881 2899 BY SIMILARITY.  
 FT DISULFID 2893 2912 BY SIMILARITY.  
 FT DISULFID 2919 2932 BY SIMILARITY.  
 FT DISULFID 2926 2945 BY SIMILARITY.  
 FT DISULFID 2939 2956 BY SIMILARITY.  
 FT DISULFID 2961 2973 BY SIMILARITY.  
 FT DISULFID 2968 2986 BY SIMILARITY.  
 FT DISULFID 2980 2997 BY SIMILARITY.  
 FT DISULFID 3006 3019 BY SIMILARITY.  
 FT DISULFID 3014 3032 BY SIMILARITY.

Query Match 61.5%; Score 32; DB 1; Length 4753;  
 Best Local Similarity 70.0%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSLRKRLTY 10  
 Db 463 KSLSLKSLVY 472

RESULT 9  
 MISS\_SCHPO STANDARD; PRT: 868 AA.  
 AC P49731;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE MISS PROTEIN.  
 GN MISS.  
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
 OC EUKARYOTA: FUNGI: ASCOMYCOTA: ARCHIAZCOMYCETES:  
 CC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;  
 CC SCHIZOSACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 95170112.  
 RA TAKAHASHI K., YAMADA H., YANAGIDA M.;  
 RT "Fission yeast minichromosome loss mutants mis cause lethal  
 RT aneuploidy and replication abnormality.";  
 RL MOL. BIOL. CELL 5:1145-1158(1994).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D31960; G829086; .  
 DR PROSITE: PS00847; MCM\_1; 1.  
 DR PROSITE: PS50051; MCM\_2; 1.  
 DR PRAM: PF00493; MCM; 1.  
 KM TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW DNA REPLICATION; CELL CYCLE; ATP-BINDING.  
 FT DOMAIN 426 633 MCM.  
 FT NP\_BIND 476 483 ATP (POTENTIAL).  
 SO SEQUENCE 868 AA; 96753 MW; EAF353B0 CRC32;

Query Match 61.5%; Score 32; DB 1; Length 868;  
 Best Local Similarity 54.5%; Pred. No. 54;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KSLSLRKRLTYR 11  
 Db 360 KSLGVRLTYK 370

RESULT 10  
 R13\_PORPU STANDARD; PRT: 126 AA.  
 AC P51295;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S13.  
 GN RPS13.  
 OS PORPHYRA PURPUREA.  
 OG CHLOROPLAST.  
 OC EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; BANGIALES; BANGIACEAE; PORPHYRA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AVONPORT.  
 RA REITH M.E., MUMHOLLAND J.;  
 RT "Complete nucleotide sequence of the porphyrin purpurea chloroplast  
 RT genome.";  
 RL PLANT MOL. BIOL. REP. 13:333-335(1995).  
 CC -1- FUNCTION: INVOLVED IN THE BINDING OF FMET-tRNA AND, HENCE, IN THE  
 CC INITIATION OF TRANSLATION (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U38804; G1276761; .  
 DR PROSITE: PS00646; RIBOSOMAL\_S13; 1.  
 DR PRAM: PF00416; S13; 1.  
 DR MENDEL; 10345; PORPU; RPS13; 1.  
 KM RIBOSOMAL PROTEIN; CHLOROPLAST.  
 SO SEQUENCE 126 AA; 14538 MW; F64E61F3 CRC32;

Query Match 61.5%; Score 32; DB 1; Length 126;  
 Best Local Similarity 53.3%; Pred. No. 6.3;  
 Matches 8; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

QY 1 KSLSLRKRL----TYR 11  
 Db 74 ESMKRLMEISTYR 88

RESULT 11  
 VGLG\_RABYT STANDARD; PRT: 524 AA.  
 AC P32550;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
 DE SPIKE GLYCOPROTEIN PRECURSOR.  
 GN G.  
 OS RABIES VIRUS (STRAIN STREET).  
 OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES;  
 CC RHABDOVIRIDAE; LYSSAVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 92142521.  
 RA BENMANSOUR A., BRAHIMI M., TUFFEREAU C., COULON P., LAFAY F.,  
 RA FLAMAND A.;  
 RT "Rapid sequence evolution of street rabies glycoprotein is related to

or send an email to license@isb-sib.ch).

-----

CC EMBL: M96150; G156360; -  
 CC EMBL: 273907; E1346098; -  
 DR WORMPEP; F29D11.1; CE05765.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 6.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 3.  
 DR PROSITE; PS01187; EGF\_CA; 3.  
 DR PROSITE; PS01209; LDLRA\_1; 27.  
 DR PROSITE; PS00068; LDLRA\_2; 34.  
 DR PRAM; PF00057; ldl\_recept\_a; 35.  
 DR PRAM; PF00058; ldl\_recept\_b; 25.  
 DR HSP; P07204; 1EGT  
 DR RECEPTOR; TRANSMEMBRANE; REPEAT; ENDOCYTOSIS; GLYCOPROTEIN;  
 KW SIGNAL; CALCIUM-BINDING; EGF-LIKE DOMAIN; COATED PITS.  
 FT FT SIGNAL 1 18 POTENTIAL.  
 FT FT CHAIN 19 4753 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN.  
 FT FT DOMAIN 19 4570 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 4571 4596 POTENTIAL.  
 FT DOMAIN 4597 4753 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 51 89 LDL-RECEPTOR CLASS A 1.  
 FT DOMAIN 90 133 LDL-RECEPTOR CLASS A 2.  
 FT DOMAIN 136 177 LDL-RECEPTOR CLASS A 3.  
 FT DOMAIN 180 220 LDL-RECEPTOR CLASS A 4.  
 FT DOMAIN 221 259 LDL-RECEPTOR CLASS A 5.  
 FT DOMAIN 260 298 LDL-RECEPTOR CLASS A 6.  
 FT DOMAIN 299 337 EGF-LIKE 1.  
 FT DOMAIN 338 368 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 369 712 EGF-LIKE 3.  
 FT DOMAIN 669 712 EGF-LIKE 4.  
 FT DOMAIN 997 1043 EGF-LIKE 5.  
 FT DOMAIN 1052 1097 LDL-RECEPTOR CLASS A 7.  
 FT DOMAIN 1099 1140 LDL-RECEPTOR CLASS A 8.  
 FT DOMAIN 1144 1184 LDL-RECEPTOR CLASS A 9.  
 FT DOMAIN 1185 1225 LDL-RECEPTOR CLASS A 10.  
 FT DOMAIN 1226 1265 LDL-RECEPTOR CLASS A 11.  
 FT DOMAIN 1268 1309 LDL-RECEPTOR CLASS A 12.  
 FT DOMAIN 1311 1352 LDL-RECEPTOR CLASS A 13.  
 FT DOMAIN 1357 1397 LDL-RECEPTOR CLASS A 14.  
 FT DOMAIN 1398 1436 EGF-LIKE 6.  
 FT DOMAIN 1437 1476 EGF-LIKE 7.  
 FT DOMAIN 1747 1786 EGF-LIKE 8.  
 FT DOMAIN 2080 2120 EGF-LIKE 9.  
 FT DOMAIN 2396 2439 EGF-LIKE 10.  
 FT DOMAIN 2728 2780 LDL-RECEPTOR CLASS A 15.  
 FT DOMAIN 2790 2831 LDL-RECEPTOR CLASS A 16.  
 FT DOMAIN 2832 2870 LDL-RECEPTOR CLASS A 17.  
 FT DOMAIN 2872 2914 LDL-RECEPTOR CLASS A 18.  
 FT DOMAIN 2917 2958 LDL-RECEPTOR CLASS A 19.  
 FT DOMAIN 2959 2999 LDL-RECEPTOR CLASS A 20.  
 FT DOMAIN 2999 3046 LDL-RECEPTOR CLASS A 21.  
 FT DOMAIN 3047 3095 LDL-RECEPTOR CLASS A 22.  
 FT DOMAIN 3098 3137 LDL-RECEPTOR CLASS A 23.  
 FT DOMAIN 3138 3176 LDL-RECEPTOR CLASS A 24.  
 FT DOMAIN 3185 3223 EGF-LIKE 11.  
 FT DOMAIN 3224 3265 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 3266 3306 EGF-LIKE 13.  
 FT DOMAIN 3582 3624 LDL-RECEPTOR CLASS A 25.  
 FT DOMAIN 3625 3668 LDL-RECEPTOR CLASS A 26.  
 FT DOMAIN 3669 3707 LDL-RECEPTOR CLASS A 27.  
 FT DOMAIN 3707 3748 LDL-RECEPTOR CLASS A 28.  
 FT DOMAIN 3751 3790 LDL-RECEPTOR CLASS A 29.  
 FT DOMAIN 3791 3832 LDL-RECEPTOR CLASS A 30.  
 FT DOMAIN 3831 3873 LDL-RECEPTOR CLASS A 31.  
 FT DOMAIN 3876 3914 LDL-RECEPTOR CLASS A 32.  
 FT DOMAIN 3915 3957 LDL-RECEPTOR CLASS A 33.  
 FT DOMAIN 3957 3997 LDL-RECEPTOR CLASS A 34.  
 FT DOMAIN 3998 4042 LDL-RECEPTOR CLASS A 35.  
 FT DOMAIN 4047 4085 EGF-LIKE 14.  
 FT DOMAIN 4088 4131 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 4132 4176 EGF-LIKE 16.  
 FT DOMAIN 4477 4515

FT DOMAIN 4526 4554 EGF-LIKE 17.  
 FT SITE 4658 4658 CRITICAL FOR ENDOCYTOSIS (BY SIMILARITY).  
 FT SITE 4744 4744 CRITICAL FOR ENDOCYTOSIS (BY SIMILARITY).  
 FT DISULFID 53 65 BY SIMILARITY.  
 FT DISULFID 60 78 BY SIMILARITY.  
 FT DISULFID 72 87 BY SIMILARITY.  
 FT DISULFID 92 109 BY SIMILARITY.  
 FT DISULFID 99 122 BY SIMILARITY.  
 FT DISULFID 116 131 BY SIMILARITY.  
 FT DISULFID 138 152 BY SIMILARITY.  
 FT DISULFID 146 165 BY SIMILARITY.  
 FT DISULFID 159 175 BY SIMILARITY.  
 FT DISULFID 182 195 BY SIMILARITY.  
 FT DISULFID 189 208 BY SIMILARITY.  
 FT DISULFID 202 218 BY SIMILARITY.  
 FT DISULFID 223 248 BY SIMILARITY.  
 FT DISULFID 230 257 BY SIMILARITY.  
 FT DISULFID 242 275 BY SIMILARITY.  
 FT DISULFID 262 288 BY SIMILARITY.  
 FT DISULFID 269 297 BY SIMILARITY.  
 FT DISULFID 282 311 BY SIMILARITY.  
 FT DISULFID 302 320 BY SIMILARITY.  
 FT DISULFID 307 336 BY SIMILARITY.  
 FT DISULFID 322 352 BY SIMILARITY.  
 FT DISULFID 342 367 BY SIMILARITY.  
 FT DISULFID 348 361 BY SIMILARITY.  
 FT DISULFID 363 367 BY SIMILARITY.  
 FT DISULFID 673 682 BY SIMILARITY.  
 FT DISULFID 678 697 BY SIMILARITY.  
 FT DISULFID 699 711 BY SIMILARITY.  
 FT DISULFID 1001 1010 BY SIMILARITY.  
 FT DISULFID 1006 1026 BY SIMILARITY.  
 FT DISULFID 1028 1042 BY SIMILARITY.  
 FT DISULFID 1054 1068 BY SIMILARITY.  
 FT DISULFID 1063 1081 BY SIMILARITY.  
 FT DISULFID 1075 1095 BY SIMILARITY.  
 FT DISULFID 1101 1114 BY SIMILARITY.  
 FT DISULFID 1108 1127 BY SIMILARITY.  
 FT DISULFID 1121 1138 BY SIMILARITY.  
 FT DISULFID 1146 1158 BY SIMILARITY.  
 FT DISULFID 1153 1171 BY SIMILARITY.  
 FT DISULFID 1165 1182 BY SIMILARITY.  
 FT DISULFID 1187 1199 BY SIMILARITY.  
 FT DISULFID 1194 1212 BY SIMILARITY.  
 FT DISULFID 1206 1223 BY SIMILARITY.  
 FT DISULFID 1228 1241 BY SIMILARITY.  
 FT DISULFID 1235 1254 BY SIMILARITY.  
 FT DISULFID 1248 1263 BY SIMILARITY.  
 FT DISULFID 1270 1283 BY SIMILARITY.  
 FT DISULFID 1277 1296 BY SIMILARITY.  
 FT DISULFID 1290 1307 BY SIMILARITY.  
 FT DISULFID 1313 1325 BY SIMILARITY.  
 FT DISULFID 1320 1338 BY SIMILARITY.  
 FT DISULFID 1332 1350 BY SIMILARITY.  
 FT DISULFID 1339 1373 BY SIMILARITY.  
 FT DISULFID 1366 1386 BY SIMILARITY.  
 FT DISULFID 1380 1396 BY SIMILARITY.  
 FT DISULFID 1401 1412 BY SIMILARITY.  
 FT DISULFID 1408 1421 BY SIMILARITY.  
 FT DISULFID 1423 1435 BY SIMILARITY.  
 FT DISULFID 1441 1451 BY SIMILARITY.  
 FT DISULFID 1447 1460 BY SIMILARITY.  
 FT DISULFID 1462 1475 BY SIMILARITY.  
 FT DISULFID 1475 1490 BY SIMILARITY.  
 FT DISULFID 1751 1760 BY SIMILARITY.  
 FT DISULFID 1756 1770 BY SIMILARITY.  
 FT DISULFID 1772 1785 BY SIMILARITY.  
 FT DISULFID 2084 2095 BY SIMILARITY.  
 FT DISULFID 2091 2105 BY SIMILARITY.  
 FT DISULFID 2107 2119 BY SIMILARITY.  
 FT DISULFID 2400 2415 BY SIMILARITY.  
 FT DISULFID 2411 2426 BY SIMILARITY.  
 FT DISULFID 2438 2438 BY SIMILARITY.  
 FT DISULFID 2732 2743 BY SIMILARITY.

```

RN [1]
RP SEQUENCE FROM N.A.
RA FARES H.F., PRINGLE J.R.
RL SUBMITTED (XXX-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 1-1553 FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE: 96310631.
RA SEM-GUPTA M., LYCK R., FLEIG U., NIEDENTHAL R.K., HEGEMANN J.H.;
RT "The sequence of a 24,152 bp segment from the left arm of chromosome
RT XIV from Saccharomyces cerevisiae between the BntI and the PstI
RT genes."
RL YEAST 12:505-514(1996).
RN [3]
RP SEQUENCE OF 988-1953 FROM N.A.
RA MESSENGER F., DUBOIS E., VIERENDELS F., SCHERENS B., PIERARD A.,
RA GLANSDORFF N.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: TO S.POMBE FUS1 AND S.POMBE CDC12.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL: L31766; G472525; -
DR EMBL: Z71546; E328804; -
DR EMBL: Z71547; E239654; -
DR SGD: X92494; G1045237; -
DR SGD: L0000190; BNT1.
FT DOMAIN 64 67 POLY-SER.
FT DOMAIN 1053 1057 POLY-SER.
FT DOMAIN 1239 1250 POLY-PRO.
FT DOMAIN 1278 1291 POLY-PRO.
FT DOMAIN 1303 1309 POLY-PRO.
FT DOMAIN 1751 1754 POLY-GLU.
FT CONFLICT 938 938 T -> A (IN REF. 1).
FT CONFLICT 1430 1430 G -> C (IN REF. 1).
SQ SEQUENCE 1953 AA; 219701 MW; B3C66ABD CRC32;

Query Match 61.5%; Score 32; DB 1; Length 1953;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSLSKRL 8
Db 964 KSLAKRL 971

RESULT 7
HS7L_SBYV STANDARD: PRT: 598 AA.
AC P37092;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE HEAT SHOCK PROTEIN 70 HOMOLOG.
OS SUGAR BEET YELLOW VIRUS (SBYV).
OC VIRUSES: SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; CLOSTEROVIRIDAE;
OC CLOSTEROVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-UKRAINIAN;
RX MEDLINE: 9116305.
RA DOJLA V.V., BOYKO V.P., KARASEV A.V., LUNINA N.A., KOONIN E.V.,
RA AGRANOVSKY A.A.;
RT "Nucleotide sequence of the 3'-terminal half of beet yellows
RT closterovirus RNA genome: unique arrangement of eight virus genes.";
RL J. GEN. VIROL. 72:15-23(1991).

```

```

RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE: 91171285.
RA AGRANOVSKY A.A., BOYKO V.P., KARASEV A.V., KOONIN E.V., DOJLA V.V.;
RT "Putative 65 kDa protein of beet yellows closterovirus is a homologue
RT of HSP70 heat shock proteins."
RL J. MOL. BIOL. 217:603-610(1991).
RN [1]
RP SEQUENCE OF 598 AA; 65230 MW; 4AAC437E CRC32;
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL: X53462; G58880; -
DR EMBL: X73476; G405628; -
DR PIR: S28712; S28712
DR PROSITE: P500297; HSP70_1; FALSE_NEG.
DR PROSITE: P500329; HSP70_2; 1.
DR PROSITE: P501036; HSP70_3; 1.
KW ATP-BINDING.
SQ SEQUENCE 598 AA; 65230 MW; 4AAC437E CRC32;

Query Match 61.5%; Score 32; DB 1; Length 598;
Best Local Similarity 69.2%; Pred. No. 36;
Matches 9; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 1 KSLSKRLT--YR 11
Db 547 KALFLKRLTADYR 559

RESULT 8
LRP_CAEEL STANDARD: PRT: 4753 AA.
ID LRP_CAEEL
AC Q04833;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN PRECURSOR (LRP).
GN LRP OR P39D11.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93281621.
RA YOCHEM J., GREENWALD I.;
RT "A gene for a low density lipoprotein receptor-related protein in the
RT nematode Caenorhabditis elegans."
RL PROC. NATL. ACADE. SCI. U.S.A. 90:4572-4576(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WILKINSON J.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: MAY ACT AS A RECEPTOR FOR THE ENDOCYTOSIS OF
CC EXTRACELLULAR LIGANDS SUCH AS CHYLOMICRON REMNANTS, PROTEASE-
CC INHIBITOR COMPLEXES AND VITELLOGENIN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS 35 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 17 EGF-LIKE REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

DR EMBL: 273320; E245584; -  
 DR EMBL: U53879; G1262308; -  
 DR PIR: S22782; S22782.  
 DR PIR: S22783; S22783.  
 DR PIR: A41943; A41943.  
 DR SGD: L0001374; PEP3.  
 DR PFAM: PF00637; Clathrin\_repeat; 1.  
 KM ZINC-FINGER; MEMBRANE.  
 FT ZN\_FING 826 851  
 SO SEQUENCE 918 AA; 107398 MW; B7A5303D CRC32;

Query Match 65.4%; Score 34; DB 1; Length 918;  
 Best Local Similarity 72.7%; Pred. No. 21;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTYR 11  
 1111111111  
 DB 354 KALSLKGLTVR 364

RESULT 4  
 ID YE29 METJA STANDARD; PRT; 503 AA.  
 AC 058824:  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN MJ1429.  
 GN MJ1429.  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 CC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE: 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLADEN A.,  
 RA OVERBECK R., KIRKNESS E.F., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA SCOTT J.L., GEORGHEN N.S.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA UTERBACK T.R., KELLEY J.M., HURST M.A., KAINE B.P., BORODOVSKY M.,  
 RA COTTON M.D., ROBERTS K.M., SMITH H.O., WOSE C.R., VENTER J.C.;  
 RA KLEIN H.-P., FRASER C.M., SMITH H.O., WOSE C.R., VENTER J.C.;  
 "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii".  
 RT JANNASCHII.  
 RL SCIENCE 273:1058-1073(1996).  
 CC -1- SIMILARITY: TO M.JANNASCHII MJ0208 AND MJ1565.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U67584; G1592079; -  
 DR TIGR: MJ1429; -  
 KM HYPOTHETICAL PROTEIN.  
 SO SEQUENCE 503 AA; 56696 MW; F8C515D1 CRC32;

Query Match 65.4%; Score 34; DB 1; Length 503;  
 Best Local Similarity 70.0%; Pred. No. 11;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTYR 10  
 1111111111  
 DB 324 KHLLEKRLTYR 333

RESULT 5  
 ID FENR\_CYAPA STANDARD; PRT; 363 AA.  
 AC 000598;  
 DT 01-JUL-1993 (REL. 26, CREATED)  
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE FERREDOXIN-NADP REDUCTASE PRECURSOR (EC 1.18.1.2) (FNR).  
 GN PETH.  
 OS CYANOPHORA PARADOXA.  
 CC EUKARYOTA; GLAUCOCYSTOPHYCEAE; CYANOPHORACEAE; CYANOPHORA.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 66-90.  
 RC STRAIN-LB555 UTEX;  
 RX MEDLINE: 93257619.  
 RA JAKOWITSCH J., BAYER M.G., NAUER T.L., LUETTKE A., GEBHART U.B.,  
 RA BRANDTNER M., HAMILTON B., NEUMANN-SPALIART C., MICHALOWSKI C.B.,  
 RA BOHNER H.J., SCHENK H.E.A., LOEFFELHARDT W.;  
 "Sequence analysis of pre-ferredoxin-NADP(+) reductase cDNA from  
 Cyanophora paradoxa specifying a precursor for a nucleus-encoded  
 cyanelle polypeptide".  
 RT PLANT MOL. BIOL. 21:1023-1033(1993).  
 RL FUNCTION: MAY PLAY A KEY ROLE IN REGULATING THE RELATIVE AMOUNTS  
 CC OF CYCLIC AND NON-CYCLIC ELECTRON FLOW TO MEET THE DEMANDS OF THE  
 CC PLANT FOR ATP AND REDUCING POWER.  
 CC -1- CATALYTIC ACTIVITY: REDUCED FERREDOXIN + NADP(+) - OXIDIZED  
 CC FERREDOXIN + NADPH.  
 CC -1- COFACTOR: FAD FLAVOPROTEIN.  
 CC -1- SIMILARITY: WITH OTHER SPECIES FNR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X66372; G18100; -  
 DR PIR: S33545; S33545.  
 DR PFAM: PF00175; oxidored\_fad; 1.  
 DR MENDEL: 7857; CYAPA; Peth; 1.  
 DR HSP: P00455; 1FNB.  
 DR OXIDOREDUCTASE: FLAVOPROTEIN; NADP; FAD; CYANELLE; TRANSIT PEPTIDE;  
 KW ELECTRON TRANSPORT; PHOTOSYNTHESIS; THYLAKOID MEMBRANE.  
 FT TRANSIT 1 65  
 FT CHAIN 66 363  
 FT DISULFD 181 186  
 FT NP BIND 215 233  
 FT SEQUENCE 363 AA; 40491 MW; 2094DB3E CRC32;

Query Match 63.5%; Score 33; DB 1; Length 363;  
 Best Local Similarity 77.8%; Pred. No. 13;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 SLSKRLTYR 10  
 1111111111  
 DB 161 SLSVKRLTYR 169

RESULT 6  
 ID BN1L\_YEAST STANDARD; PRT; 1953 AA.  
 AC P41832;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE BN1L PROTEIN (SYNTHETIC LETHAL 39).  
 GN BN1L OR SHES OR YNL271C OR N0646.  
 OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 CC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
 CC SACCCHAROMYCETACEAE; SACCCHAROMYCES.



50 SEQUENCE 366 AA: 39203 MW: D9067EE3 CRC32:

Query Match Best Local Similarity 84.6%: Score 44: DB 1: Length 366;

Matches 10: Conservative 0: Mismatches 1: Indels 0: Gaps 0;

QY 1 KSLSLKRLTYR 11  
| | | | | | | |  
Db 83 KSLSLKRLTYR 93

RESULT 2

ID CD14\_RAT STANDARD: PRT: 371 AA.

AC 063691:

DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE MONOCYTE DIFFERENTIATION ANTIGEN CD14 PRECURSOR (LPS RECEPTOR) (LPS-R)

DE (WEILOID CELL-SPECIFIC LECINE-RICH GLYCOPROTEIN).

GN CD14.

OS RATTUS NORVEGICUS (RAT).

OC EURAROTIA; METAZOA: CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.

[1]

SEQUENCE FROM N.A.

RC TISSUE-SPLEEN.

RC MEDLINE: 96175065.

RA GALEA E., REIS D.J., FOX E.S., XU H., FEINSTEIN D.L.;

RT "CD14 mediate endotoxin induction of nitric oxide synthase in

RT cultured brain glial cells."

RL J. NEUROIMMUNOL. 64:19-28(1996).

[2]

SEQUENCE FROM N.A.

RA MEDLINE: 97344898.

RA TATAI N., KATAOKA M., HIGUCHI Y., MATSURA K., YAMAMOTO S.;

RT "Primary structure of rat CD14 and characteristics of rat CD14,

RT cytokine, and NO synthase mRNA expression in mononuclear phagocyte

RT system cells in response to LPS."

RL J. LEUKOC. BIOL. 61:736-744(1997).

[3]

SEQUENCE FROM N.A.

AP TISSUE-LIVER:

RC MEDLINE: 99003114.

RA LIU S., KHEMLANI L.S., SHAPIRO R.A., JOHNSON M.L., LIU K.,

RA GELLER D.A., WATKINS S.C., GOWERT S.M., BILLIAR T.R.;

RT "Expression of CD14 by hepatocytes: upregulation by cytokines during

RT endotoxemia."

RL INFECT. IMMUN. 66:5089-5098(1998).

-1- FUNCTION: SERVES AS AN LPS RECEPTOR CONTROLLING CELL ACTIVATION

UNDER PHYSIOLOGICAL CONDITIONS. WHEN LPS BINDS TO CD14 THE CELLS

BECOME ACTIVATED AND RELEASE CYTOKINES AND UP-REGULATE CELL

SURFACE MOLECULES, INCLUDING ADHESION MOLECULES (BY SIMILARITY).

-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U51804; G1262820; -

DR EMBL: AF087943; G3599525; -

DR PFAM: PF00560; LRR: 3.

KW ANTIGEN; GLYCOPROTEIN; GPI-ANCHOR; SIGNAL.

FT CHAIN 1 17 POTENTIAL.

FT CARBOHYD 18 371 CD14 DIFFERENTIATION ANTIGEN.

FT CARBOHYD 39 39 POTENTIAL.

FT CARBOHYD 46 46 POTENTIAL.

FT CARBOHYD 152 152 POTENTIAL.

FT CARBOHYD 185 185 POTENTIAL.

FT CARBOHYD 281 281 POTENTIAL.

FT CONFLICT 55 55 G -> GA (IN REF. 3).

FT CONFLICT 219 219 S -> P (IN REF. 3).

FT CONFLICT 268 268 A -> T (IN REF. 3).

SQ SEQUENCE 371 AA: 39942 MW: 40339386 CRC32:

Query Match Best Local Similarity 69.2%: Score 36: DB 1: Length 371;

Matches 8: Conservative 1: Mismatches 2: Indels 0: Gaps 0;

QY 1 KSLSLKRLTYR 11  
| | | | | | | |  
Db 88 KSLSLKRLTYR 98

RESULT 3

ID PEP3\_YEAST STANDARD: PRT: 918 AA.

AC P27801:

DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)

DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)

DE VACUOLAR MEMBRANE PROTEIN PEP3.

GN PEP3 OR VPS18 OR YL148W OR L9634.2.

OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EURAROTIA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;

OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.

[1]

SEQUENCE FROM N.A.

RC MEDLINE: 92049305.

RA PRESTON R., MANOLSON M.F., BECHERER K., WEINDHAMMER E.;

RT "Isolation and characterization of PEP3, a gene required for vacuolar

RT biogenesis in *Saccharomyces cerevisiae*."

RL MOL. CELL. BIOL. 11:5801-5812(1991).

[2]

SEQUENCE FROM N.A.

RC MEDLINE: 92049306.

RA ROBINSON J.S., GRAHAM T.R., EMR S.D.;

RT "A putative zinc finger protein, *Saccharomyces cerevisiae* Vps18p,

RT affects late Golgi functions required for vacuolar protein sorting

RT and efficient alpha-factor prohormone maturation."

RL MOL. CELL. BIOL. 11:5813-5824(1991).

[3]

SEQUENCE FROM N.A.

RA RIEGER M., MUELLER-AUER S., BRUECKNER M.;

RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

[4]

SEQUENCE FROM N.A.

RC STRAIN-S288C / AB972;

RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.;

RA FAVELLO A., FULON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.;

RA HALLSMORF K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.;

RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENZES S.;

RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIKKEN L., RILES L.;

RA TATCH A., TREVASKIS E., VIGNATI D., WILCOX L., WOLDMAN P., VAUDIN M.;

RA WILSON R., WATERSTON R.;

RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

-1- FUNCTION: REQUIRED FOR VACUOLAR BIOGENESIS.

-1- SUBCELLULAR LOCATION: CYTOSOLIC FACE OF THE VACUOLAR MEMBRANE.

-1- SIMILARITY: SOME, TO DROSOPHILA DEEP ORANGE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M65144; G172120; -



**THIS PAGE BLANK (USPTO)**

A:Reference number: 215023  
 A:Accession: T03716  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-87 <HR>  
 A:Cross-references: EMBL:D85877; NID:d1069785; PID:d1013588  
 A:Experimental source: subsp. japonica, cv. Nipponbare

Query Match 61.5%; Score 32; DB 2; Length 87;  
 Best Local Similarity 63.6%; Pred. No. 7.4;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KSLSLKLRLYR 11  
 ||| ||: ||  
 Db 33 KSLSLKLRLYR 43

## RESULT 15

S63244  
 BNI1 protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein N0646; protein YNL271c  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 27-Apr-1996 #sequence revision 03-May-1996 #text\_change 17-Mar-1999  
 C:Accession: S63244; S63245; S48523; S60909; S65111  
 R:Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.  
 submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S63235  
 A:Accession: S63244  
 A:Molecule type: DNA  
 A:Residues: 1-1553 <SEN>  
 A:Cross-references: EMBL:Z71547; MIPS:YNL271c  
 R:Experimental source: strain S288C  
 R:Mesenguy, F.; Dubois, E.; Verendeels, F.; Scherens, B.; Pierard, A.; Glansdorff, N.  
 submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S63245  
 A:Accession: S63245  
 A:Molecule type: DNA  
 A:Residues: 987-1953 <MES>  
 A:Cross-references: EMBL:Z71547; MIPS:YNL271c  
 R:Experimental source: strain S288C  
 R:Fares, H.F.; Pringle, J.R.  
 submitted to the EMBL Data Library, April 1994  
 A:Description: Synthetic lethals of CDC12.  
 A:Reference number: S48523  
 A:Accession: S48523  
 A:Molecule type: DNA  
 A:Residues: 1-937, 'A', 939-1429, 'C', 1431-1953 <FAR>  
 A:Cross-references: EMBL:L31766; NID:g472524; PID:g472525  
 R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.  
 submitted to the EMBL Data Library, October 1995  
 A:Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV fr  
 A:Reference number: S60909  
 A:Accession: S60909  
 A:Molecule type: DNA  
 A:Residues: 1-1553 <SE2>  
 A:Cross-references: EMBL:X92494; NID:g1045236; PID:g1045237  
 R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.  
 yeast 12, 505-514, 1996  
 A:Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa  
 A:Reference number: S65111; M01D:96310631  
 A:Accession: S65111  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1553 <SEN>  
 A:Cross-references: EMBL:X92494; NID:g1045236; PID:g1045237  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
 C:Genetics:  
 A:Gene: SGD:BNI1; SHE5; SYL39  
 A:Cross-references: SGD:S0005215; MIPS:YNL271c  
 A:Map position: 14L

Query Match 61.5%; Score 32; DB 2; Length 1953;  
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSLSLKLRL 8  
 |||:||||  
 Db 964 KSLSLKLRL 971

Search completed: September 7, 1999, 23:22:35  
 Job time: 942 sec

Query Match 61.5% Score 32; DB 2; Length 4753;  
 Best Local Similarity 70.0%; Pred. No. 5.4e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KSLSLKRLTY 10  
 |||||  
 Db 463 KKLKSLKSLVY 472

RESULT 10  
 S73216  
 Ribosomal protein S13, chloroplast - red alga (*Porphyra purpurea*) chloroplast  
 C:Species: chloroplast *Porphyra purpurea*  
 C:Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 30-Jan-1998  
 C:Accession: S73216  
 R:Reith, M.; Munholland, J.  
 Plant Mol. Biol. Rep. 13, 333-335, 1995  
 A:Title: Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome.  
 A:Reference number: S73108  
 A:Accession: S73216  
 A:Molecule type: DNA  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Residues: 1-126 <REI>  
 A:Cross-references: EMBL:U38804; NID:g127652; PID:g1276761  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
 C:Genetics:  
 A:Gene: rps13  
 A:Genome: chloroplast  
 A:Start codon: GTG  
 C:Superfamily: Escherichia coli ribosomal protein S13  
 C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 61.5% Score 32; DB 2; Length 126;  
 Best Local Similarity 53.3%; Pred. No. 11;  
 Matches 8; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

OY 1 KSLSLKRL---TYR 11  
 :||:||||  
 Db 74 ESMISIKRLMEISTR 88

RESULT 11  
 S28712  
 heat shock protein 70 homolog - sugar beet yellows virus  
 C:Species: sugar beet yellows virus, SBVY  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 24-Sep-1998  
 C:Accession: S28712  
 R:Agarovsky, A.A.; Boyko, V.P.; Karasev, A.V.; Lunina, N.A.; Koonin, E.V.; Dolja, V.V.  
 J. Gen. Virol. 72, 15-23, 1991  
 A:Title: Nucleotide sequence of the 3'-terminal half of beet yellows closterovirus RNA  
 A:Reference number: S28710  
 A:Accession: S28712  
 A:Molecule type: DNA  
 A:Residues: 1-598 <AGR>  
 A:Cross-references: EMBL:X53462; NID:g58878; PID:g58880  
 C:Superfamily: sugar beet yellows virus heat shock protein 70 homolog

Query Match 61.5% Score 32; DB 2; Length 598;  
 Best Local Similarity 69.2%; Pred. No. 58;  
 Matches 9; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 1 KSLSLKRLTY--YR 11  
 ||:|||||  
 Db 547 KALFLKRLTADR 559

RESULT 12  
 G69753  
 two-component sensor histidine kinase homolog ycbm - *Bacillus subtilis*  
 C:Species: *Bacillus subtilis*  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998

C:Accession: G69753  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
 C: Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
 A.: Ehlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gal  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
 Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scant  
 A:Authors: Schleich, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Se  
 akuch, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*  
 A:Reference number: A69580; MUID:98044033  
 A:Accession: G69753  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-158 <RUN>  
 A:Cross-references: GB:299105; GB:AL009126; NID:g2632457; PID:el182208; PID:g2632542  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: ycbm

Query Match 61.5% Score 32; DB 2; Length 158;  
 Best Local Similarity 55.6%; Pred. No. 14;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 SLSLKRLTY 10  
 :||:||||  
 Db 150 TITLKRMTY 158

RESULT 13  
 S72834  
 hypothetical protein B1620.F2.68 - *Mycobacterium leprae*  
 C:Species: *Mycobacterium leprae*  
 C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Sep-1997  
 C:Accession: S72834  
 R:Smith, D.R.; Robison, K.  
 submitted to the EMBL Data Library, November 1993  
 A:Description: *Mycobacterium leprae* cosmid B1620.  
 A:Reference number: S72580  
 A:Accession: S72834  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-577 <SMI>  
 A:Cross-references: EMBL:U00015; NID:g466931; PID:g466955  
 C:Genetics:  
 A:Start codon: GTG

Query Match 61.5% Score 32; DB 2; Length 577;  
 Best Local Similarity 75.0%; Pred. No. 56;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSLKRLTY 10  
 ||:|||||  
 Db 107 LSLRLTY 114

RESULT 14  
 T03716  
 reverse transcriptase homolog - rice retrotransposon Tos18 (fragment)  
 C:Species: *Oryza sativa* (rice)  
 C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
 C:Accession: T03716  
 R:Hiroyuki, H.; Sugimoto, K.; Otsuki, Y.; Tsugawa, H.; Kanda, M.  
 Proc. Natl. Acad. Sci. U.S.A. 93, 7783-7788, 1996  
 A:Title: Retrotransposons of rice involved in mutations induced by tissue culture.

F:70-210/Domain: FAD binding #status predicted <FAD>  
 F:94-348/Domain: cytochrome-b5 reductase homology <CBR>  
 F:211-363/Domain: NADP binding #status predicted <NADP>

Query Match 63.5%; Score 33; DB 2; Length 363;  
 Best Local Similarity 77.8%; Pred. No. 21;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 SLSLKRLTY 10  
 |||:||||  
 Db 161 SLSVKRLTY 169

RESULT 6  
 D69211  
 conserved hypothetical protein MTH835 - Methanobacterium thermoautotrophicum (strain Del  
 C:Species: Methanobacterium thermoautotrophicum  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
 C:Accession: D69211  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Delouche, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
 R:Qu, D.; Spadafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivani, N.  
 K: S. Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
 A:Reference number: A69000; MUID:98037514  
 A:Accession: D69211  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-324 <MTH>  
 A:Cross-references: GB:AE000860; GB:AE000666; NID:g2621915; PID:g2621925  
 A:Experimental source: strain Delta H  
 A:Genetics:  
 A:Gene: MTH835  
 A:Start codon: TTG

Query Match 63.5%; Score 33; DB 2; Length 324;  
 Best Local Similarity 63.6%; Pred. No. 18;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTYR 11  
 ::|:|||||  
 Db 112 RTCSDKRLTYR 122

RESULT 7  
 S50216  
 translation initiation factor eIF-2 alpha chain kinase (EC 2.7.1.-) - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 14-Jul-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
 C:Accession: S50216  
 R:Mellor, H.; Flowers, K.M.; Kimball, S.R.; Jefferson, L.S.  
 Biochim. Biophys. Acta 1219, 693-696, 1994  
 A>Title: Cloning and characterization of a cDNA encoding rat PKR, the double-stranded RN  
 A:Reference number: S50216; MUID:95035100  
 A:Accession: S50216  
 A:Status: preliminary; translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-513 <MEL>  
 A:Cross-references: EMBL:L29281; NID:g468372; PID:g468373  
 C:Superfamily: protein kinase TIK; double-stranded RNA-binding repeat homology; protein  
 C:Keywords: autophosphorylation; phosphoprotein; phosphotransferase; RNA binding; serine  
 F:5-77/Domain: double-stranded RNA-binding repeat homology <DSR>  
 F:234-500/Domain: protein kinase homology <KIN>

Query Match 61.5%; Score 32; DB 2; Length 513;  
 Best Local Similarity 50.0%; Pred. No. 49;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 KSLSLKRLTY 10  
 |: :|:|:|

Db 260 KTYAIKRLTY 269

RESULT 8  
 I56563  
 interleukin-3 receptor beta-subunit - rat  
 C:Species: Rattus sp. (rat)  
 C>Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 16-Dec-1998  
 C:Accession: I56563  
 R:Appel, K.; Buttini, M.; Sauter, A.; Gebicke-Haerter, P.J.  
 J. Neurosci. 15, 5800-5809, 1995  
 A>Title: Cloning of rat interleukin-3 receptor beta-subunit from cultured microglia a  
 A:Reference number: I56563; MUID:95370942  
 A:Accession: I56563  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-896 <RES>  
 A:Cross-references: GB:S79263; NID:g1086954; PID:g1086955  
 C:Genetics:  
 A:Gene: rll-3b beta  
 C:Superfamily: interleukin-3 receptor beta chain; cytokine receptor homology  
 C:Keywords: cytokine receptor  
 F:39-235/Domain: cytokine receptor homology <CRS1>  
 F:253-433/Domain: cytokine receptor homology <CRS2>

Query Match 61.5%; Score 32; DB 2; Length 896;  
 Best Local Similarity 54.5%; Pred. No. 90;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTYR 11  
 |: :|:|:|:|  
 Db 852 KDLVSKRLPYQ 862

RESULT 9  
 A47437  
 LDL-receptor-related protein - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 24-Sep-1998  
 C:Accession: A47437; S27801  
 R:Yochem, J.; Greenwald, I.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 4572-4576, 1993  
 A>Title: A gene for a low density lipoprotein receptor-related protein in the nematod  
 A:Reference number: A47437; MUID:93281621  
 A:Accession: A47437  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-4753 <YOC>  
 A:Cross-references: GB:M96150; NID:g156359; PID:g156360  
 A:Note: nucleotide sequence not given; translation not complete in this paper  
 R:Yochem, J.; Greenwald, I.  
 submitted to the EMBL Data Library, July 1992  
 A:Description: A gene for an LDL receptor-related protein (LPR) in the nematode C. ele  
 A:Reference number: S27801  
 A:Accession: S27801  
 A:Molecule type: DNA  
 A:Residues: 1-4753 <YOC>  
 A:Cross-references: EMBL:M96150; NID:g156359; PID:g156360  
 C:Genetics:  
 A:Gene: LPR  
 A:Introns: 31/1; 88/1; 132/1; 172/3; 219/1; 298/1; 463/2; 526/2; 585/3; 780/2; 874/2;  
 15/1  
 C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-bind  
 C:Keywords: tandem repeat; transmembrane protein  
 F:1441-1475/Domain: EGF homology <EGF>  
 F:1611-1654/Domain: LDL receptor WYTD-containing repeat homology <YW33>  
 F:3187-3222/Domain: LDL receptor WYTD-containing repeat homology <LDLO>  
 F:3709-3745/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
 F:3753-3788/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
 F:4343-4386/Domain: LDL receptor WYTD-containing repeat homology <YW38>

A:Reference number: A64300; MUID:96337999  
A:Accession: D64478  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-503 <BUL>  
A:Cross-references: GB:U67564; GB:L77117; NID:g1592077; PID:g1592079; TIGR:MU1429; PID:g1592079  
A:Map position: FOR1399336-1400847  
A:Start codon: GTG  
C:Superfamily: conserved hypothetical protein MU1429

Query Match 65.4%; Score 34; DB 2; Length 503;  
Best Local Similarity 70.0%; Pred. No. 18;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTYR 10  
| | | | | | | | | |  
Db 324 KHLKRLTYR 333

RESULT 3  
C71414  
hypothetical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: Columbia  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 05-Dec-1998  
C:Accession: C71414  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk  
P.; Wedler, H.; Wedler, E.; Wandt, R.; Woltzenegger, T.; Pohl, T.M.; Terry, N.; Giel  
avanagh, T.; Hempel, S.; Kotter, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgomech  
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palmer, K.; Benes, V.; Reichen, S.; Ans  
C.; Chawatzky, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal  
A:Reference number: A71400; MUID:98121113  
A:Accession: C71414  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-555 <BEV>  
A:Cross-references: GB:297337; NID:g2244829; PID:e326846; PID:g2244863  
C:Genetics:  
A:Map position: 4COP9-463845

Query Match 65.4%; Score 34; DB 2; Length 555;  
Best Local Similarity 63.6%; Pred. No. 20;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTYR 11  
| | | | | | | | | |  
Db 360 KHLKRLTYR 370

RESULT 4  
A1943  
vacuolar membrane protein PEP3 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein L313; protein YLR148w  
C:Species: Saccharomyces cerevisiae  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Feb-1998  
C:Accession: A1943; S22782; S64997; S22783  
R:Robinson, J.S.; Graham, T.R.; Emt, S.D.  
Mol. Cell. Biol. 11, 5813-5824, 1991  
A:Title: A putative zinc finger protein, Saccharomyces cerevisiae Vps18p, affects late c  
A:Reference number: A1943; MUID:92049306  
A:Accession: A1943  
A:Molecule type: DNA  
A:Residues: 1-918 <ROB>  
A:Cross-references: EMBL:M65144; NID:g172119; PID:g172120  
A:Note: sequence extracted from NCBI backbone (NCBIN:65599, NCBI:65601)  
R:Preston, R.A.; Manolson, M.F.; Becherer, K.; Weidenhammer, E.; Kirkpatrick, D.; Wright  
Mol. Cell. Biol. 11, 5801-5812, 1991

A:Title: Isolation and characterization of PEP3, a gene required for vacuolar biogenesis  
A:Reference number: S22782; MUID:92049305  
A:Accession: S22782  
A:Molecule type: DNA  
A:Residues: 1-918 <PRE>  
A:Cross-references: EMBL:M65144; NID:g172119; PID:g172120  
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64987  
A:Accession: S64997  
A:Molecule type: DNA  
A:Residues: 1-918 <RIE>  
A:Cross-references: EMBL:Z73320; NID:g1360575; PID:e245584; PID:g1360576; MIPS:YLR148  
A:Note: experimental\_source strain S288C  
C:Genetics:  
A:Gene: SCD:PEP3; VPS18; VPT18; VAM8  
A:Cross-references: SGD:S0004138; MIPS:YLR148w  
A:Map position: 12R  
C:Function:  
A:Description: involved in vacuolar protein sorting and required for vacuole biogenesis  
C:Keywords: transmembrane protein; yeast vacuole  
F:511-527/Domain: transmembrane #status predicted <TM>

Query Match 65.4%; Score 34; DB 2; Length 918;  
Best Local Similarity 72.7%; Pred. No. 35;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTYR 11  
| | | | | | | | | |  
Db 354 KSLSLKRLTYR 364

RESULT 5  
A5664  
ferredoxin--NADP+ reductase (EC 1.18.1.2) precursor - Cyanophora paradoxa  
C:Species: Cyanophora paradoxa  
C:Date: 31-Dec-1993 #sequence\_revision 21-Jul-1995 #text\_change 13-Sep-1998  
C:Accession: S33545; S38805; A5664; S30032  
R:Jakowitsch, J.; Bayer, M.G.; Mayer, T.L.; Luettke, A.; Gebhart, U.B.; Brandtner, M.  
Plant Mol. Biol. 21, 1023-1033, 1993  
A:Title: Sequence analysis of pre-ferredoxin--NADP(+) reductase cDNA from Cyanophora p  
A:Reference number: S33545  
A:Accession: S33545  
A:Molecule type: mRNA  
A:Residues: 1-363 <JAK>  
A:Cross-references: EMBL:X66372; NID:g18099; PID:g18100  
A:Experimental source: strain LB555 UTEX  
A:Accession: S38805  
A:Molecule type: protein  
A:Residues: 66-90 <JAK>  
A:Experimental source: strain LB555 UTEX  
R:Gebhart, U.B.; Mayer, T.L.; Stevanovic, S.; Bayer, M.G.; Schenk, H.E.  
Protein Expr. Purif. 3, 228-235, 1992  
A:Title: Ferredoxin:NADP oxidoreductase of Cyanophora paradoxa: purification, partial  
A:Reference number: A5664; MUID:93005754  
A:Accession: A5664  
A:Molecule type: protein  
A:Residues: 67-90 <GBB>  
A:Note: a shorter form lacking the first four residues from the amino end was also fo  
A:Note: sequence extracted from NCBI backbone (NCBI:119664)  
C:Genetics:  
A:Gene: peth  
A:Genome: nuclear  
A:Complex: monomer  
C:Function:  
A:Description: catalyzes electron transfer from reduced ferredoxin to NADP  
A:Pathway: photosynthesis  
A:Note: FAD cofactor  
C:Superfamily: ferredoxin--NADP+ reductase; cytochrome-b5 reductase homology  
C:Keywords: cyanelle; electron transfer; FAD; flavoprotein; NADP; oxidoreductase; pho  
F:1-65/Domain: transit peptide (cyanelle) #status experimental <TMP>  
F:66-363/Product: ferredoxin--NADP+ reductase #status experimental <MAT>

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:32; Search time 105.14 Seconds  
(without alignments)  
4.192 Million cell updates/sec

Title: US-09-124-280A-22  
Perfect score: 52  
Sequence: 1 KSLSLKRLTYR 11

Scoring table: BIOSUM62

Searched: 122810 seqs, 40065486 residues

Database :  
1: PIR.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	84.6	366	1	TDMSM4	monocyte surface g
2	34	65.4	503	2	D64478	hypothetical prote
3	34	65.4	555	2	C71414	hypothetical prote
4	34	65.4	918	2	A41943	vacuolar membrane
5	33	63.5	363	2	A56664	ferredoxin--NADP+
6	33	63.5	324	2	D59211	conserved hypothet
7	32	61.5	513	2	S50216	translactin initia
8	32	61.5	896	2	I56563	interleukin-3 rece
9	32	61.5	4753	2	A47437	LDL-receptor-relat
10	32	61.5	126	2	S73216	ribosomal protein
11	32	61.5	126	2	S28712	heat shock protein
12	32	61.5	158	2	G69753	two-component sens
13	32	61.5	577	2	S72834	hypothetical prote
14	32	61.5	87	2	T03716	reverse transcript
15	32	61.5	1953	2	S63244	BN1 protein - yea
16	31	59.6	2767	1	UIHU	thyroglobulin prec
17	31	59.6	2769	1	UIBO	thyroglobulin
18	31	59.6	101	1	WSULRB	E5 protein - cotto
19	31	59.6	512	1	ERADN4	early E2A DNA-bind
20	31	59.6	448	2	G69960	exodeoxyribonuclea
21	31	59.6	675	2	PQ0227	adenylate cyclase
22	31	59.6	1251	2	S48687	type VIII adenylyl
23	31	59.6	1248	2	A53588	adenylate cyclase
24	31	59.6	127	2	S06347	thyroglobulin - sh
25	31	59.6	450	2	JQ1594	tubulin alpha chai
26	31	59.6	50	2	S41325	8K protein - potat
27	31	59.6	70	2	S03197	8K protein - potat
28	31	59.6	1038	2	S37854	hypothetical prote
29	31	59.6	537	2	H69503	medium-chain acyl-
30	31	59.6	431	2	G71874	hypothetical prote
31	31	59.6	413	2	C64641	nickel and cobalt
32	31	59.6	717	2	S77190	hypothetical prote
33	31	59.6	99	2	JH0205	DNA-(apurinic or a
34	31	59.6	536	2	T02937	probable kinase as
35	31	59.6	585	2	S88098	hypothetical prote
36	31	59.6	83	2	D71608	hypothetical prote
37	31	59.6	244	2	S45361	UKR47 protein - fr
38	31	59.6	428	2	S17448	probable ligand-bi
39	31	59.6	473	2	S17448	

40	31	59.6	309	2	E71170	hypothetical prote
41	30	57.7	292	1	A26952	electron transfer
42	30	57.7	1024	1	GBEC	beta-galactosidase
43	30	57.7	121	1	R3B513	ribosomal protein
44	30	57.7	1115	1	IJMSNL	neural cell adhesi
45	30	57.7	569	2	S31214	hypothetical prote

## ALIGNMENTS

```

RESULT 1
TDMSM4
monocyte surface glycoprotein CD14 precursor - mouse
N:Alternate names: myeloid cell-specific leucine-rich glycoprotein
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 14-Nov-1997
C:Accession: S03605; A43539; S04967
R:Matsunura, K.; Setoguchi, M.; Nasu, N.; Higuchi, Y.; Yoshida, S.; Yamam
Nucleic Acids Res. 17, 2132, 1989
A:Title: Nucleotide and amino acid sequences of the mouse CD14 gene.
A:Reference number: S03605; MUID:89183627
A:Accession: S03605
A:Molecule type: DNA
A:Residues: 1-366 <MATS>
A:Cross-references: EMBL:X13987; NID:g192460; PID:g387121
R:Setoguchi, M.; Nasu, N.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.
Biochim. Biophys. Acta 1008, 213-222, 1989
A:Title: Mouse and human CD14 (myeloid cell-specific leucine-rich glycoprotein) prima
J. Immunol. 145, 331-336, 1990
A:Title: CD14 is a member of the family of leucine-rich proteins and is encoded by a
A:Reference number: A43539; MUID:90293480
A:Accession: A43539
A:Molecule type: DNA
A:Residues: 1-366 <FER>
A:Cross-references: GB:M34510; NID:g192460; PID:g387121
R:Setoguchi, M.; Nasu, N.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.
Biochim. Biophys. Acta 1008, 213-222, 1989
A:Title: Mouse and human CD14 (myeloid cell-specific leucine-rich glycoprotein) prima
A:Reference number: S04967; MUID:89287330
A:Accession: S04967
A:Molecule type: mRNA
A:Residues: 1-366 <SET>
A:Cross-references: EMBL:X13333; NID:g50338; PID:g50339
C:Genetics: 1/3
A:Insertions: 1/3
C:Superfamily: monocyte surface glycoprotein CD14
C:Keywords: glycoprotein
F:1-15/Domain: signal sequence #status predicted <Sig>
F:16-366/Product: monocyte surface glycoprotein CD14 #status predicted <MAT>
F:87-323/Region: 9-residue repeats (LIV)-X-X-L-X-[LVIY]-[SRAL]-X-[NAT])
F:33,147,180,276,317/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 84.6%; Score 44; DB 1; Length 366;
Best Local Similarity 90.9%; Pred. No. 0.096;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 KSLSLKRLTYR 11
DB 83 KSLSLKRLTYR 93
RESULT 2
D64478
hypothetical protein MJ1429 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Sep-1998
C:Accession: D64478
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc

```



**THIS PAGE BLANK (USPTO)**

CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09235  
FILING DATE: Concurrently  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-193  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 472 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS: LINEAR  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
PCT-US94-09235-2

Query Match 55.8%; Score 29; DB 3; Length 472;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KSLSLKRLT 9  
||| :||  
Db 178 KSLSCORIT 186

Search completed: September 7, 1999, 23:07:44  
Job time: 1742 sec

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 870 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-732-192A-2

Query Match 55.8%; Score 29; DB 2; Length 870;  
Best Local Similarity 45.5%; Pred. No. 3.3e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLRRLTYR 11  
| : : : : : |  
Db 478 KNTLRQLVYR 488

RESULT 13  
US-08-804-227C-9  
; Sequence 9, Application US/08804227C  
; Patent No. 5876991  
; GENERAL INFORMATION:  
; APPLICANT: Dehoff, Bradley S.  
; APPLICANT: Kuhstoss, Stuart A.  
; APPLICANT: Rostreck, Paul R., Jr.  
; APPLICANT: Sutton, Kimberly L.  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THOMAS G. PLANT 1501  
; STREET: LILLY CORPORATE CENTER  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII(DOS) Text only  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/804,227C  
; FILING DATE: February 21, 1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plant, Thomas, G.  
; REGISTRATION NUMBER: 35,784  
; REFERENCE/DOCKET NUMBER: X-8231  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-2459  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1996 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-804-227C-9

Query Match 55.8%; Score 29; DB 2; Length 1996;  
Best Local Similarity 55.6%; Pred. No. 7.4e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSLSLRRLTY 9  
| : : : : : |  
Db 1255 RSLALRLTY 1263

RESULT 14  
US-08-337-358-37  
; Sequence 37, Application US/08337358  
; Patent No. 5879676

; GENERAL INFORMATION:  
; APPLICANT: Liu, Chi-Li  
; APPLICANT: Adams, Lee F.  
; APPLICANT: Luburrow, Patricia A.  
; APPLICANT: Thomas, Michael D.  
; TITLE OF INVENTION: NOVEL BACILLUS THURINGIENSIS STRAINS  
; TITLE OF INVENTION: ACTIVE AGAINST LEPIDOPTERAN AND COLEOPTERAN PESTS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 58796760 No. 5879676disk of No. 5879676ch America, Inc.  
; STREET: 405 Lexington Avenue, 64th floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Tape  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/337,358  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/264,100  
; FILING DATE: 22-JUN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/194,651  
; FILING DATE: 09-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/166,391  
; FILING DATE: 13-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/991,073  
; FILING DATE: 15-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arris Dr., Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 3778.230-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 287 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-337-358-37

Query Match 55.8%; Score 29; DB 2; Length 287;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLRRLTY 10  
| : : : : : |  
Db 22 KGYDLKRVSY 31

RESULT 15  
PCT-US94-09235-2  
; Sequence 2, Application PC/TUS9409235  
; GENERAL INFORMATION:  
; APPLICANT: LI, ET AL.  
; TITLE OF INVENTION: Calcitonin Receptor  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD

NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,865  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1403 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-865-17

Query Match 57.7%; Score 30; DB 2; Length 1403;  
Best Local Similarity 54.5%; Pred. No. 3.4e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
DB 1227 KTLFLSKRTYR 1237

RESULT 11  
US-08-433-783-37  
Sequence 37, Application US/08433783  
Patent No. 5770431  
GENERAL INFORMATION:  
APPLICANT: Liu, Chi-Li  
APPLICANT: Adams, Lee F.  
APPLICANT: Lufburrow, Patricia A.  
APPLICANT: Thomas, Michael D.  
TITLE OF INVENTION: NOVEL BACILLUS THURINGIENSIS STRAINS  
TITLE OF INVENTION: ACTIVE AGAINST LEPIDOPTERAN AND COLEOPTERAN PESTS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 57704310 No. 5770431disk of No. 5770431th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Tape  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433,783  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,358

FILING DATE: 10-NOV-1994  
APPLICATION NUMBER: US 08/264,100  
FILING DATE: 22-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/194,651  
FILING DATE: 09-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/166,391  
FILING DATE: 13-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/991,073  
FILING DATE: 15-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Agtis Dr., Cheryl H.  
REGISTRATION NUMBER: 34,086  
REFERENCE/DOCKET NUMBER: 3778.230-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 287 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-433-783-37

Query Match 55.8%; Score 29; DB 2; Length 287;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSLSLKRLTY 10  
DB 22 KGVYDKRVSX 31

RESULT 12  
US-08-732-192A-2  
Sequence 2, Application US/08732192A  
Patent No. 5849526  
GENERAL INFORMATION:  
APPLICANT: Pichersky, Eran  
TITLE OF INVENTION: USE OF LINALOOL SYNTHASE IN GENETIC  
TITLE OF INVENTION: ENGINEERING OF SCENT PRODUCTION  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE & DURKEE  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/732,192A  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,146  
FILING DATE: 25-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: HIGHLANDER, STEVEN L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UMIC-015  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/005,002C  
FILING DATE: 15-JAN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/789,915  
FILING DATE: 08-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Farrell, Kevin M.  
REGISTRATION NUMBER: 35,505  
REFERENCE/DOCKET NUMBER: MIT-5091A3Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 207-363-0558  
TELEFAX: 207-363-0528  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1121 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-005-002C-2

Query Match 57.7% Score 30; DB 1; Length 1121;  
Best Local Similarity 54.5% Pred. No. 2.7e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
I:I:I I I I I  
DB 945 KTLFISRTYR 955

RESULT 8  
US-07-908-253-3  
Sequence 3, Application US/07908253  
Patent No. 5534256  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: HARLAND, RICHARD J.  
TITLE OF INVENTION: HAEMOPHILUS SOMNUS OUTER MEMBRANE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ROBERTA L. ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/908,253  
FILING DATE: 19920702  
CLASSIFICATION: 420  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9000-0026  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1403 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-908-253-3

Query Match 57.7% Score 30; DB 1; Length 1403;  
Best Local Similarity 54.5% Pred. No. 3.4e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
I:I:I I I I I  
DB 1227 KTLFISRTYR 1237

RESULT 9  
US-08-487-203A-2  
Sequence 2, Application US/08487203A  
Patent No. 5683904  
GENERAL INFORMATION:  
APPLICANT: Baker, Rohan T.  
APPLICANT: Tobias, John W.  
APPLICANT: Varshavsky, Alexander  
TITLE OF INVENTION: Ubiquitin-Specific Proteases  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kevin M. Farrell, P.C.  
STREET: P.O. Box 999  
CITY: York Harbor  
STATE: Maine  
COUNTRY: U.S.A.  
ZIP: 03911  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,203A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/005,002  
FILING DATE: 15-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Farrell, Kevin M.  
REGISTRATION NUMBER: 35,505  
REFERENCE/DOCKET NUMBER: MIT-5091A3Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 207-363-0558  
TELEFAX: 207-363-0528  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1121 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-487-203A-2

Query Match 57.7% Score 30; DB 1; Length 1121;  
Best Local Similarity 54.5% Pred. No. 2.7e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
I:I:I I I I I  
DB 945 KTLFISRTYR 955

RESULT 10  
US-08-694-865-17  
Sequence 17, Application US/08694865  
Patent No. 5837268  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

RESULT 5  
US-07-853-913-2  
; Sequence 2, Application US/07853913  
; Patent No. 5338839  
GENERAL INFORMATION:  
APPLICANT: McKay, Ronald D.G.  
APPLICANT: Lendahl, Urban  
TITLE OF INVENTION: Nestin Expression As An Indicator of  
Neuroepithelial Tumors  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/853,913  
FILING DATE: 19920319  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/660,412  
FILING DATE: 22-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,803  
FILING DATE: 25-OCT-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/201,762  
FILING DATE: 02-JUN-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/180,548  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: MIT-4641AAAA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1805 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-853-913-2

Query Match 57.7%; Score 30; DB 1; Length 1805;  
Best Local Similarity 70.0%; Pred. No. 4.3e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 SLSLKRLTYR 11  
|||||  
DB 295 SLSLEVATYR 304

RESULT 6  
US-07-853-913-4  
; Sequence 4, Application US/07853913  
; Patent No. 5338839  
GENERAL INFORMATION:  
APPLICANT: McKay, Ronald D.G.  
APPLICANT: Lendahl, Urban  
TITLE OF INVENTION: Nestin Expression As An Indicator of  
Neuroepithelial Tumors  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/853,913  
FILING DATE: 19920319  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/660,412  
FILING DATE: 22-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,803  
FILING DATE: 25-OCT-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/201,762  
FILING DATE: 02-JUN-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/180,548  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: MIT-4641AAAA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1618 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-853-913-4

Query Match 57.7%; Score 30; DB 1; Length 1618;  
Best Local Similarity 70.0%; Pred. No. 3.9e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 SLSLKRLTYR 11  
|||||  
DB 293 SLSLEVATYR 302

RESULT 7  
US-08-005-002C-2  
; Sequence 2, Application US/08005002C  
; Patent No. 5494818  
GENERAL INFORMATION:  
APPLICANT: Baker, Rohan T.  
APPLICANT: Tobias, John W.  
APPLICANT: Varshavsky, Alexander  
TITLE OF INVENTION: Ubiquitin-Specific Proteases  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kevin M. Farrell, P.C.  
STREET: P.O. Box 999  
CITY: York Harbor  
STATE: Maine  
COUNTRY: U.S.A.  
ZIP: 03911  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8998  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-22

Query Match 100.0%; Score 52; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSLSKRLTYR 11  
|||||  
Db 1 KSLSKRLTYR 11

RESULT 3  
PCT-US96-08295-40  
Sequence 40, Application PC/US9608295  
GENERAL INFORMATION:  
APPLICANT: Massachusetts Institute of Technology  
TITLE OF INVENTION: REGULATORS OF G-PROTEIN SIGNALING  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/08295  
FILING DATE: 31-MAY-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/588,258  
FILING DATE: 12-JAN-96  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bleker-Brady, Kristina  
REGISTRATION NUMBER: 39,109  
REFERENCE/DOCKET NUMBER: 01997/216001  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-08295-40

Query Match 59.6%; Score 31; DB 3; Length 420;  
Best Local Similarity 77.8%; Pred. No. 66;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSLSKRLT 9  
|||: |||  
Db 406 KSLSKRLT 414

RESULT 4  
US-07-789-915A-2  
Sequence 2, Application US/07789915A  
Patent No. 5212058  
GENERAL INFORMATION:  
APPLICANT: Baker, Rohan T.  
APPLICANT: Varshavsky, Alexander  
TITLE OF INVENTION: Ubiquitin-Specific Proteases  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millita Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/789,915A  
FILING DATE: 19911108  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: MIT-5091AA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1121 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-789-915A-2

Query Match 57.7%; Score 30; DB 1; Length 1121;  
Best Local Similarity 54.5%; Pred. No. 2.7e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSKRLTYR 11  
|||: |||  
Db 945 KTLFISKRLTYR 955

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:43 ; Search time 82.37 Seconds  
(without alignments)  
1.318 Million cell updates/sec

Title: US-09-124-280a-22  
Perfect score: 52  
Sequence: 1 KSLSLKRLTYR 11

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database: Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/PCTUS9.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	52	100.0	11	1	US-08-097-830E-22	Sequence 22, Appl
2	52	100.0	11	2	US-08-456-112B-22	Sequence 22, Appl
3	31	59.6	420	3	PCT-US96-08295-40	Sequence 40, Appl
4	30	57.7	1121	1	US-07-789-915A-2	Sequence 2, Appl
5	30	57.7	1805	1	US-07-853-913-2	Sequence 2, Appl
6	30	57.7	1618	1	US-07-853-913-4	Sequence 4, Appl
7	30	57.7	1121	1	US-08-005-002C-2	Sequence 2, Appl
8	30	57.7	1403	1	US-07-908-253-3	Sequence 3, Appl
9	30	57.7	1121	1	US-08-487-203A-2	Sequence 2, Appl
10	30	57.7	1403	2	US-08-694-865-17	Sequence 17, Appl
11	29	55.8	287	2	US-08-433-783-37	Sequence 37, Appl
12	29	55.8	670	2	US-08-732-192A-2	Sequence 2, Appl
13	29	55.8	1996	2	US-08-804-227C-9	Sequence 9, Appl
14	29	55.8	287	2	US-08-337-358-37	Sequence 37, Appl
15	29	55.8	472	3	PCT-US94-09235-2	Sequence 37, Appl
16	29	55.8	287	3	PCT-US95-07537A-37	Sequence 37, Appl
17	29	55.8	287	3	PCT-US95-07537A-37	Sequence 37, Appl
18	28	53.8	27	1	US-08-090-036-10	Sequence 10, Appl
19	28	53.8	41	1	US-08-462-949-38	Sequence 38, Appl
20	28	53.8	41	1	US-08-023-764B-38	Sequence 38, Appl
21	28	53.8	524	1	US-08-461-837-2	Sequence 2, Appl
22	28	53.8	215	1	US-08-431-080-22	Sequence 22, Appl
23	28	53.8	186	1	US-08-565-386-7	Sequence 7, Appl
24	28	53.8	149	2	US-08-602-010A-18	Sequence 18, Appl
25	28	53.8	1190	2	US-08-337-690A-2	Sequence 2, Appl
26	28	53.8	579	2	US-08-448-196A-8	Sequence 8, Appl
27	28	53.8	149	2	US-08-680-726A-18	Sequence 18, Appl
28	28	53.8	119	2	US-08-680-726A-62	Sequence 62, Appl
29	28	53.8	524	2	US-08-480-736-12	Sequence 12, Appl
30	28	53.8	524	3	PCT-US96-09495-2	Sequence 2, Appl
31	27	51.9	867	1	US-07-867-106-3	Sequence 3, Appl
32	27	51.9	921	1	US-07-872-644-39	Sequence 39, Appl
33	27	51.9	942	1	US-07-872-644-43	Sequence 43, Appl
34	27	51.9	941	1	US-07-872-644-45	Sequence 45, Appl
35	27	51.9	892	1	US-07-977-434-12	Sequence 12, Appl
36	27	51.9	473	1	US-08-103-739B-2	Sequence 2, Appl
37	27	51.9	1531	1	US-08-141-893-2	Sequence 2, Appl
38	27	51.9	402	1	US-08-236-311-1	Sequence 1, Appl
39	27	51.9	434	1	US-08-236-311-4	Sequence 4, Appl

40	27	51.9	193	1	US-08-106-507-2	Sequence 2, Appl
41	27	51.9	216	1	US-08-106-507-10	Sequence 10, Appl
42	27	51.9	921	1	US-08-297-494-39	Sequence 39, Appl
43	27	51.9	942	1	US-08-297-494-43	Sequence 43, Appl
44	27	51.9	941	1	US-08-297-494-45	Sequence 45, Appl
45	27	51.9	797	3	PCT-US95-10245-2	Sequence 2, Appl

## ALIGNMENTS

```
RESULT 1
US-08-097-830E-22
; Sequence 22, Application US/08097830E
; Patent No. 5652211
GENERAL INFORMATION:
APPLICANT: PORIO, Massimo
TITLE OF INVENTION: Peptides For Neutralizing The
TITLE OF INVENTION: Toxicity of Lipid A
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: Hedman, Gibson & Costigan, P.C.
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097/830E
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Costigan, James V.
REGISTRATION NUMBER: 25,669
TELEPHONE: (212) 302-8989
TELECOMMUNICATION INFORMATION:
TELEFAX: (212) 302-8998
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-097-830E-22

Query Match      100.0%  Score 52:  DB 1:  Length 11:
Best Local Similarity 100.0%  Pred No. 0.00016:
Matches 11:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

Oy      1 KSLSLKRLTYR 11
Db      1 KSLSLKRLTYR 11

RESULT 2
US-08-456-112B-22
; Sequence 22, Application US/08456112B
; Patent No. 5834430
GENERAL INFORMATION:
APPLICANT: PORIO, Massimo
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: Hedman, Gibson & Costigan
```



**THIS PAGE BLANK (USPTO)**

Best Local Similarity 70.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches

3; Indels 0; Gaps 0;

OY 1 KSLSLKRLTY 10  
| | | | | | |  
Db 94 KPLSLNRLQY 103

Search completed: September 7, 1999, 22:49:54  
Job time: 7875 sec

PT Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter  
PT phosphate metabolism, modulate inflammatory function or innate  
PT immunity responses  
PS Claim 3; Page 147-149; 171pp; English.  
CC The present invention specifically describes human DNAX toll-like  
CC receptors 2 to 10 (DTLR2-10). The present sequence is human DTLR4  
CC given in the present invention. Also described are: (1) a fusion  
CC protein comprising a DTLR protein or peptide; (2) a binding compound,  
CC preferably an antibody or antibody fragment which specifically binds to  
CC a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or  
CC peptide; (4) an expression vector comprising the nucleic acid of (3);  
CC and (5) a host cell comprising the vector of (4). The host cell of (5)  
CC can be used to produce the DTLR proteins. The DTLR proteins can be used  
CC to alter phosphate metabolism, to modulate inflammatory function, innate  
CC immunity responses or morphological effects. The DTLR proteins can be  
CC used in the treatment of conditions exhibiting abnormal expression of  
CC the receptors of their ligands. These abnormalities are typically  
CC manifested by immunological disorders.  
SQ Sequence 837 AA;

Query Match 59.6%; Score 31; DB 1; Length 837;  
Best Local Similarity 70.0%; Pred. No. 2.4e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTY 10  
| | | | | | | | | |  
DB 347 KLKSLKRLTF 356

## RESULT 14

W88439  
ID W88439 standard; Protein: 428 AA.

AC 26-APR-1999 (first entry)  
DE Arabidopsis pyruvate dehydrogenase subunit E1 alpha.  
KW Pyruvate dehydrogenase E1 alpha; chloroplast; transgenic plant;  
KW plasmid: poly(3-hydroxybutyrate-co-3-hydroxyvalerate) copolymer;  
KW P3HB-co-3HV; polyhydroxyalkanoate; biodegradable plastic.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT Peptide 1..68  
FT /note= "targeting peptide"  
FT Protein 69..428  
FT Binding\_site 160..213  
FT /note= "thiamine pyrophosphate binding site"  
FT Binding\_site 226..277  
FT /note= "PDH beta binding site"  
FT Modified\_site 263  
FT /note= "conserved phosphorylation site"  
FT Modified\_site 325  
FT /note= "conserved phosphorylation site"  
FT Modified\_site 332  
FT /note= "non-conserved phosphorylation site"  
PN W09900505-A1.  
PD 07-JAN-1989  
PF 02-MAR-1998; U13406  
PR 30-JUN-1998; US-076554.  
PR 30-JUN-1997; US-051291.  
PR 01-AUG-1997; US-055255.  
PR 02-MAR-1998; US-076544.  
PA (UMOR ) UNIT MISSOURI.  
PI Johnston ML, Luechly MH, Miernyk JA, Mooney BP, Randall DR;  
DR WPI: 99-095750/08.  
DR N-PSDB: X06837.  
CC New nucleic acid encoding subunits of plasmid pyruvate dehydrogenase  
CC - or branched chain 2-oxoacid dehydrogenase used for producing  
PT polyhydroxyalkanoate polymers in high yield  
PS Claim 4; Page 91-92; 151pp; English.  
CC This polypeptide comprises the E1 alpha subunit of the Arabidopsis  
CC thaliana plasmid pyruvate dehydrogenase (PDH) complex. The  
CC invention provides nucleotide sequences that encode the E1 alpha

CC and E1 beta subunits, and the and E2 component (see X06837-39) of  
CC the plasmid PDH complex, as well as the E1 alpha and E1 beta  
CC subunits and E2 component (see X06840-42) of the branched chain  
CC 2-oxoacid dehydrogenase complex of A. thaliana. These nucleotide  
CC sequences, and the encoded proteins (see W88439-44), can be  
CC introduced into plants in various combinations in order to enhance  
CC the conversion of threonine to 2-oxobutyrate, propionate,  
CC propionyl-CoA, beta-ketovalleryl-CoA and beta-hydroxyvaleryl-CoA.  
CC Introduction into such plants of nucleic acid sequences encoding an  
CC appropriate beta-keto-thiolase, a beta-ketoacyl-CoA reductase and a  
CC polyhydroxyalkanoate synthase will permit such transgenic plants to  
CC utilise the increased beta-hydroxyvaleryl-CoA substrate in the  
CC production of poly(3-hydroxybutyrate-co-3-hydroxyvalerate)  
CC copolymer, a biodegradable plastic used to make mouldings, films,  
CC coatings and in drug-release applications.  
SQ Sequence 428 AA;

Query Match 59.6%; Score 31; DB 1; Length 428;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KSLSLKRLTY 10  
: | | | | | | | | | |  
DB 41 RSLSLRLNH 50

## RESULT 15

Y12477  
ID Y12477 standard; Protein: 131 AA.

AC Y12477;  
DT 17-JUN-1999 (first entry)

DE Human 5' EST secreted protein SEQ ID NO:508.  
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;  
KW forensic; gene therapy; chromosome mapping; signal peptide;  
KW upstream regulatory sequence; cytokine activity; cell proliferation;  
KW differentiation; haemopoiesis regulation; tissue growth regulation;  
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
KW thrombolytic; anti-inflammatory; tumour inhibition.  
OS Homo sapiens.  
PN W09906548-A2.  
PN 11-FEB-1999.  
PF 31-JUL-1998; IB1222.  
PR 01-AUG-1997; US-905135.  
PA (GENSET ) GENSET.  
PI Duclert A, Dumas Malne Edwards J, Lacroix B;  
DR WPI: 99-153778/13.  
DR N-PSDB: X41310.  
CC New nucleic acids encoding human secreted proteins - obtained from  
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,  
PT kidney, lung, umbilical cord, placenta and colon tissue  
PS Claim 27; Page 798-799; 824pp; English.  
CC X41094 to X41347 represent 5' expressed sequence tags (ESTs) for human  
CC secreted proteins, and encode the proteins given in Y12261 to Y12514,  
CC respectively. The proteins given represent the signal peptide and an  
CC N-terminal fragment of a secreted protein. The nucleic acid sequences  
CC can be used for producing secreted human gene products. They can also  
CC be used to develop products for diagnosis and therapy. The proteins  
CC obtained may have cytokine activity, cell proliferation/differentiation  
CC activity, haematopoiesis regulating activity, tissue growth regulating  
CC activity, reproductive hormone regulating activity, chemotactic/  
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/  
CC ligand activity, anti-inflammatory activity, tumour inhibition activity  
CC or other activities. The products can be used in forensic, gene therapy  
CC and chromosome mapping procedures. The sequences can also be used for  
CC obtaining corresponding promoter sequences. The nucleic acids encoding  
CC the signal peptide can be used for directing extracellular secretion of  
CC a polypeptide or the insertion of a polypeptide into a membrane, or  
CC importing a polypeptide into a cell.  
SQ Sequence 131 AA;

Query Match 59.6%; Score 31; DB 1; Length 131;

CC or adhesion formation, heal wounds and prevent restenosis following  
 CC balloon angioplasty. Modulation of HA production in vivo may be  
 CC useful in, e.g. Graves disease, mesothelioma, Wilms' tumour and  
 CC oedema associated with inflammation of lung and kidney, all of  
 CC which are associated with elevated levels of HA in tissues or  
 CC serum.

SQ Sequence 552 AA;

Query Match 59.6%; Score 31; DB 1; Length 552;

Best Local Similarity 85.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLKRLTY 10

Db 105 SVKRLTY 111.

RESULT 11

ID W6768 standard; Protein; 241 AA.

AC W6768;

DT 15-JAN-1999 (first entry)

DE Human HOECFII protein.

KW HOECFII; hyaluronan synthase; diagnosis; detection; disease; imbalance;

KW gene therapy; immune response; vaccine; chronic renal failure; cancer;

KW inflammatory diseases; myocardial ischemia; rheumatoid arthritis; human;

KW cirrhotic liver disease; chromosome mapping; linkage analysis.

OS Homo sapiens.

PN EP-881294-A2.

PD 02-DEC-1998.

PF 20-MAY-1998; 303991.

PR 29-MAY-1997; US-865273.

PA (SMIK ) SMITHLINE BEECHAM CORP.

PI Namhi P, Pullen M, Zhu Y;

DR WPI: 99-001400/01.

N-PSDB: V62994.

PT New hyaluronan synthase (HOECFII) polypeptide and polynucleotide -

PT useful as diagnostic reagents and for prevention and treatment of

PT cancer and chronic renal failure

PS Claim 1; Page 16; 18PP; English.

CC This sequence represents a human hyaluronan synthase (HOECFII). HOECFII

CC polypeptides and polynucleotides are useful for diagnosing susceptibility

CC to diseases by detecting mutations in the HOECFII gene using probes

CC containing the HOECFII nucleotide sequence, and can diagnose diseases

CC associated with HOECFII imbalance by determining HOECFII polypeptide

CC levels. HOECFII polypeptides can be used to screen for agonists and

CC antagonists which bind HOECFII polypeptide by observing the binding, or

CC stimulation or inhibition of HOECFII activity. These can be used in

CC treatment to activate (agonist) or inhibit (antagonist) HOECFII activity,

CC in addition to direct administration of antisense sequences to prevent

CC expression, or HOECFII polypeptides to treat conditions associated with a

CC lack of HOECFII protein. Gene therapy may also be used to affect

CC endogenous HOECFII polypeptide production. HOECFII antibodies are useful

CC for inducing an immune response to immunise and prevent diseases, and for

CC isolating HOECFII clones or purifying HOECFII polypeptides by affinity

CC chromatography. HOECFII polypeptides can be administered directly or as a

CC vaccine to inoculate against disease. Diseases diagnosed, prevented or

CC treated include chronic renal failure, inflammatory diseases, myocardial

CC ischemia, cancer, rheumatoid arthritis and cirrhotic liver disease. The

CC HOECFII polypeptide is also useful for mapping the gene to a chromosome,

CC allowing gene inheritance to be studied through linkage analysis.

SQ Sequence 241 AA;

Query Match 59.6%; Score 31; DB 1; Length 241;

Best Local Similarity 85.7%; Pred. No. 66;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLKRLTY 10

Db 105 SVKRLTY 111

RESULT 12

ID W86352 standard; Protein; 799 AA.

AC W86352;

DT 15-MAR-1999 (first entry)

DE Human DNAX toll-like receptor DTLR4.

KW DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;

KW interleukin 1 receptor; phosphate metabolism; innate immunity response;

KW modulate inflammatory function; morphological effect;

KW immunological disorder.

OS Homo sapiens.

PN WO9850547-A2.

PD 12-NOV-1998.

PF 07-MAR-1998; 008979.

PR 05-MAR-1998; US-076947.

PR 07-MAY-1997; US-044293.

PR 22-JAN-1998; US-072212.

PA (SCHE ) SCHERING CORP.

PI Bazan JF, Hardiman GT, Kastelein RA, Rock FL;

DR WPI: 99-059670/05.

N-PSDB: V80665.

PT Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter

PT phosphate metabolism, modulate inflammatory function or innate

PT immunity responses

PS Example; Page 115-117; 171PP; English.

CC The present invention specifically describes human DNAX toll-like

CC receptors 2 to 10 (DTLR2-10). The present sequence is human DTLR4

CC given in the present invention. Also described are: (1) a fusion protein

CC comprising a DTLR protein or peptide; (2) a binding compound, preferably

CC an antibody or antibody fragment which specifically binds to a DTLR

CC protein or peptide; (3) a nucleic acid encoding a DTLR protein or

CC peptide; (4) an expression vector comprising the nucleic acid of (3);

CC and (5) a host cell comprising the vector of (4). The host cell of (5)

CC can be used to produce the DTLR proteins. The DTLR proteins can be used

CC to alter phosphate metabolism, to modulate inflammatory function, innate

CC immunity responses or morphological effects. The DTLR proteins can be

CC used in the treatment of conditions exhibiting abnormal expression of

CC the receptors of their ligands. These abnormalities are typically

CC manifested by immunological disorders.

SQ Sequence 799 AA;

Query Match 59.6%; Score 31; DB 1; Length 799;

Best Local Similarity 70.0%; Pred. No. 2.3e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSLSLKRLTY 10

Db 309 KKLKSLKRLTF 318

RESULT 13

ID W86361 standard; Protein; 837 AA.

AC W86361;

DT 15-MAR-1999 (first entry)

DE Human DNAX toll-like receptor DTLR4.

KW DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;

KW interleukin 1 receptor; phosphate metabolism; innate immunity response;

KW modulate inflammatory function; morphological effect;

KW immunological disorder.

OS Homo sapiens.

PN WO9850547-A2.

PD 12-NOV-1998.

PF 07-MAR-1998; 008979.

PR 05-MAR-1998; US-076947.

PR 07-MAY-1997; US-044293.

PR 22-JAN-1998; US-072212.

PA (SCHE ) SCHERING CORP.

PI Bazan JF, Hardiman GT, Kastelein RA, Rock FL;

DR WPI: 99-059670/05.

N-PSDB: V80675.

FT Modified\_site /note= "N-glycosylated"  
 FT 345  
 FT Modified\_site /note= "N-glycosylated"  
 FT 377  
 FT Modified\_site /note= "N-glycosylated"  
 FT 529  
 FT Modified\_site /note= "N-glycosylated"  
 FT 843  
 FT Modified\_site /note= "N-glycosylated"  
 FT 865  
 FT Modified\_site /note= "N-glycosylated"  
 FT 904  
 FT Modified\_site /note= "N-glycosylated"  
 FT 1076  
 FT Modified\_site /note= "N-glycosylated"  
 FT 1103  
 FT Modified\_site /note= "N-glycosylated"  
 FT 1225  
 FT Modified\_site /note= "N-glycosylated"  
 FT 1320  
 FT Modified\_site /note= "N-glycosylated"  
 FT 1392  
 FT Modified\_site /note= "N-glycosylated"  
 FT 1593  
 FT Modified\_site /note= "N-glycosylated"  
 FT 1626  
 FT Modified\_site /note= "N-glycosylated"  
 FT 1703  
 FT Modified\_site /note= "O-phosphorylated"  
 FT 1719  
 FT Modified\_site /note= "O-phosphorylated"  
 PN W09745449-A1.  
 PD 04-DEC-1997.  
 PR 29-MAY-1997; NZ00068.  
 PR 29-MAY-1996; NZ-286692.  
 PA (HART/) HART D N J.  
 PI Hart DMJ;  
 DR WPI: 98-032580/03.  
 DR N-PSDB: V04024.  
 PT Human dendritic cell receptor, DEC-205 - useful to develop ligands for isolation and targeted cell delivery of antigen or toxin  
 PS Claim 2; Fig 11; 53pp; English.  
 CC This protein comprises the human homologue of mouse dendritic cell receptor DEC-205. Its amino acid sequence was deduced from DEC-205 cDNA (see V04024) obtained from Hodgkin's disease-derived L428 cells. It shows 77% overall identity with the mouse protein. The predicted mol.wt. is 198-205 kDa. The invention provides isolated and polynucleotides encoding these polypeptides as well as vectors which include such polynucleotides, and a method of producing recombinant DEC-205 polypeptides in host cells transformed or transfected with such vectors. Further provided are ligands (preferably an antibody or antibody binding fragment) that bind to human DEC-205, as well as constructs for use in prophylaxis and therapy that comprise such a ligand, human DEC-205 or an extracellular domain coupled to an antigen capable of inducing a protective immune response in a patient or to a toxin (e.g. ritin A chain to specifically destroy dendritic cells as part of an immunosuppressive process).  
 CC Sequence 1722 AA;

Query Match 61.5%; Score 32; DB 1; Length 1722;  
 Best Local Similarity 70.0%; Pred. No. 3.3e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSKRLTY 10  
 11 11111  
 DB 1516 KSKKLSRLTY 1525

RESULT 9  
 W10168

ID W10168 standard; Protein; 420 AA.  
 AC W10168;  
 DT 31-MAR-1997 (first entry)  
 DE Human regulator of G-protein signalling RGS2.  
 KW Regulator of G-protein signalling; RGS; RGS2; nematode;  
 KW signal transduction; transgenic animal; diagnosis; therapy;  
 KW diabetes; hyperplasia; psychiatric disorder; cardiovascular disease;  
 KW McCune-Albright syndrome; Albright hereditary osteopathy.  
 OS Homo sapiens.  
 PN W09638462-A1.  
 PD 05-DEC-1996.  
 PR 31-MAY-1996; U08295.  
 PR 02-JUN-1995; US-460505.  
 PR 12-JAN-1996; US-588258.  
 PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.  
 PI Horvitz HR, Koelle M;  
 DR WPI: 97-034298/03.  
 DR N-PSDB: T50911.  
 PT New isolated regulator of G-protein signalling genes - used to develop prods. for the diagnosis and treatment of G-protein related diseases and disorders e.g. diabetes, cardiovascular disease, etc  
 PS Example C; Page 66-67; 96pp; English.  
 CC Human RGS2 (W10168) is a member of a new family of proteins (see also W10166-77) involved in the control of heterotrimeric G-protein mediated effects in both mammalian and non-mammalian cells. The proteins have been named RGS (Regulators of G-protein signalling).  
 CC RGS2 shows homology throughout its length to nematode EGL-10 (W10167). It is encoded by a gene (T50911) old. by searching a database of expressed sequence tags. RGS polypeptides can be produced in transformed host cells for use in regulating G-protein signalling and screening of cpds. for regulatory activity. They can also be used to increase secretion of polypeptides into culture media, and in the diagnosis and treatment of G-protein related disorders.  
 CC Sequence 420 AA;

Query Match 59.6%; Score 31; DB 1; Length 420;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSLSKRLTY 9  
 11 11111  
 DB 406 KSLSKRLTY 414

RESULT 10  
 W50009  
 ID W50009 standard; Protein; 552 AA.  
 AC W50009;  
 DT 11-JUN-1998 (first entry)  
 DE Murine hyaluronan synthase-2.  
 KW Murine; hyaluronan synthase-2; HS-2; hyaluronan;  
 KW hyaluronic acid.  
 OS Mus sp.  
 PN W09800551-A2.  
 PD 08-JAN-1998.  
 PR 03-JUL-1997; U11761.  
 PR 05-MAR-1997; US-812008.  
 PR 03-JUL-1996; US-675499.  
 PA (MAYO-) MAYO FOUNDATION.  
 PI Augustine ML, McDonald JA, Spicer AP;  
 DR WPI: 98-086976/08.  
 DR N-PSDB: V18821.  
 PT Nucleic acid encoding hyaluronan synthase-2 or -3 or their active fragments - used for treating arthritis, in wound repair, for delivering drugs etc.  
 PS Claim 29; Pages 60-61; 115pp; English.  
 CC The present sequence is murine hyaluronan synthase-2 (HS-2), which can be used to alter hyaluronan, i.e. hyaluronic acid (HA), production, or for in vitro synthesis of HA.  
 CC HA can be used to treat arthritis or perforated ear drums, protect eyes during surgery, deliver drugs, prevent post-operative scarring

CC antisera. The proteins are highly immunogenic and induce a long-lasting  
 CC immune response that persists even after antimicrobial treatment. In  
 CC antibody detection assays, on sera, plasma, urine, saliva etc., they are  
 CC highly sensitive and specific. The specification also describes 69  
 CC previously unrecognised immunogenic cluster families. H. pylori antigens  
 CC are used to detect H. pylori eradication of infection, and in vaccines to  
 CC infection or to confirm eradication of infection, and in vaccines to  
 CC protect against H. pylori infection and related diseases (gastritis,  
 CC peptic ulcer, gastric adenocarcinoma/lymphoma).  
 SQ Sequence 499 AA;

Query Match 69.2%; Score 36; DB 1; Length 499;  
 Best Local Similarity 63.6%; Pred. No. 14;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
 DB 42 KSLSLKRLTYR 52

## RESULT 6

ID P90113 standard; protein: 524 AA.  
 AC P90113;  
 DT 1-NOV-1989 (first entry)  
 DE Rabies G protein  
 KM Rabies virus: glycoprotein; G protein; Nishigahara strain;  
 KW N.H. strain; vaccines; diagnosis; monoclonal antibody.  
 OS Rabies virus  
 PN EP-324157-A.  
 PD 19-JUL-1989.  
 PF 23-DEC-1988; 121602.  
 PR 26-DEC-1987; JP-330896.  
 PA (KAGA) Chemo-Sero-Therap.  
 PI Sakamoto S, Ide T, Tokiyoshi S, Yamamoto M;  
 DR WPI: 89-208179/29.  
 DR N-PSDB: N90295.  
 PT Gene fragment coding for rabies virus glyco-protein  
 PT - derived from Nishigahara strain, useful for  
 PT preparing vaccine or diagnostic reagents.  
 PS Claim 1: fig 1: 21pp: English.  
 CC Rabies G protein (see corresp. N90295). Used to react with  
 CC monoclonal antibody which neutralises virus, so  
 CC it is used in vaccines and diagnostic reagents. Obtd. from a  
 CC Nishigahara strain that has been passaged through a rabbit for >2000  
 CC passages.  
 SQ Sequence 524 AA;

Query Match 63.5%; Score 33; DB 1; Length 524;  
 Best Local Similarity 60.0%; Pred. No. 59;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSLSLKRLTY 10  
 DB 313 KSVSFRRLSY 322

## RESULT 7

ID R05077 standard; protein: 524 AA.  
 AC R05077;  
 DT 10-MAR-1993 (revised)  
 DT 11-JUL-1990 (first entry)  
 DE Salt restriction product of plasmid pUC-RNSL1 encoding rabies  
 DE preglycoprotein G protein.  
 KM Rabies: pUC-RNSL1; pre-glycoprotein G; ds.  
 OS Synthetic.  
 PN WC9000191-A.  
 PD 11-JAN-1990.  
 PF 27-JUN-1989; U02809.  
 PR 01-JUL-1988; US-214085.

PA (SAK) Salt Inst for Biol Stud.

PI Tschopp JF, Dixon DB;

DR WPI: 90-037127/05.

DR N-PSDB: Q03216.

PT Prod. of rabies glyco-protein G -

PT by culturing P.pastoris contg. DNA encoding preprotein with

PT methanol as carbon source.

PT disclosure: Table 1: 82pp: English.

CC Isolating the rabies preglycoprotein G and associating it with a promoter

CC segment of a methylophilic yeast strain and a terminator from a second

CC methylophilic strain, allows protein to be selectively expressed in a

CC P. pastoris culture (pref. ATCC 20880).

CC The inactive virus protein is useful in a vaccine as a prophylactic

CC measure for at risk individuals and animals.

SQ Sequence 524 AA;

Query Match 63.5%; Score 33; DB 1; Length 524;  
 Best Local Similarity 60.0%; Pred. No. 59;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSLSLKRLTY 10  
 DB 313 KSVSFRRLSY 322

## RESULT 8

ID W38429 standard; protein: 1722 AA.  
 AC W38429;  
 DT 21-MAY-1998 (first entry)  
 DE Human dendritic cell receptor DEC-205.  
 KW Dendritic cell receptor; DEC-205; human; ligand; cell targeting;  
 KW antigen; toxin; immunosuppressive; therapy; prophylaxis.  
 OS Homo sapiens.  
 FH Key  
 FH Peptide  
 FT Location/Qualifiers  
 FT 1..26  
 FT /label= Sig-peptide  
 FT 27..1666  
 FT /note="extracellular domain"  
 FT 27..158  
 FT /note="cysteine-rich domain"  
 FT 71..106  
 FT /note="Claim 1"  
 FT 158..212  
 FT /note="fibronectin type II domain"  
 FT 216..342  
 FT /note="carbohydrate recognition domain CRD-1"  
 FT 361..488  
 FT /note="carbohydrate recognition domain CRD-2"  
 FT 501..625  
 FT /note="carbohydrate recognition domain CRD-3"  
 FT 643..792  
 FT /note="carbohydrate recognition domain CRD-4"  
 FT 811..932  
 FT /note="carbohydrate recognition domain CRD-5"  
 FT 951..1091  
 FT /note="carbohydrate recognition domain CRD-6"  
 FT 1097..1222  
 FT /note="carbohydrate recognition domain CRD-7"  
 FT 1208..1323  
 FT /note="Claim 1"  
 FT 1239..1382  
 FT /note="carbohydrate recognition domain CRD-8"  
 FT 1394..1513  
 FT /note="carbohydrate recognition domain CRD-9"  
 FT 1539..1661  
 FT /note="carbohydrate recognition domain CRD-10"  
 FT 1667..1691  
 FT /note="transmembrane domain"  
 FT 1692..1722  
 FT /note="cytoplasmic domain"  
 FT 135  
 FT Modified\_site

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA:

Query Match 100.0%; Score 52; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSLSLRKRYR 11  
 Db 1 KSLSLRKRYR 11

RESULT 3  
 ID W60855 standard; Protein: 366 AA.  
 AC W60855;  
 DT 01-OCT-1998 (first entry)  
 DE Mouse CD14 protein.  
 KM CD14; B cell activator; bovine lactation-associated immunotropic protein;  
 KM LAIR; T cell immunodeficiency; X-linked hyper-IgM syndrome; allergy;  
 KM common variable immunodeficiency; X-linked agammaglobulinemia; vaccine;  
 KM infant feeding formulae; mouse.  
 OS Mus sp.  
 PN W0982580-A2.  
 PD 28-MAY-1998.  
 PE 18-NOV-1997; CA0880.  
 PR 18-NOV-1996; US-746883.  
 PA (WELL-) WELLESLEY HOSPITAL FOUND.  
 PI Alizadeh-Khavi K, Filipp D, Julius MH;  
 DR N-PSDB; V37229.  
 DR New bovine polypeptide that activates mammalian B cell(s) - used  
 PT e.g. to treat T cell immunodeficiency or allergy, as vaccine  
 PT adjuvant, as T cell surrogate for infants, and for monoclonal  
 PT antibody production, also specific antibodies for treating B cell  
 PT hyperactivity  
 PS Claim 14; Fig 7; 64pp; English.  
 CC This sequence is the mouse CD14 protein of the invention. The CD14  
 CC protein was used to isolate the bovine CD14 of the invention,  
 CC which is able to activate mammalian B cells. The protein is also known  
 CC as bovine lactation-associated immunotropic protein (LAIR), and is used  
 CC to activate B cells, particularly in humans. Particularly it is  
 CC administered to subjects: (a) with CD40 negative or deficient B cells;  
 CC (b) suffering from T cell immunodeficiency (e.g. X-linked hyper-IgM  
 CC syndrome, common variable immunodeficiency or X-linked  
 CC agammaglobulinemia) or allergy (i.e. with CD40 ligand negative or  
 CC defective T cells); or (c) to induce growth and differentiation of  
 CC B cells to highly productive Ig secreting cells. Particular applications  
 CC are in infant feeding formulae (as immunostimulant) and as adjuvant in  
 CC vaccines (optionally with bovine CD14 coupled to the antigen). The DNA  
 CC sequences are also used to enrich mammalian B cells secreting a  
 CC monoclonal antibody (MAb) of particular antigenic specificity, by  
 CC activating cells with sub-optimal amount of the DNA in combination with  
 CC the antigen. The enriched B cells are then used to produce hybridomas  
 CC that produce specific MAb. Antibodies raised against human CD14 are used  
 CC to reduce/inhibit activity of B cells that are hyperactivated by high  
 CC serum levels of CD14. Bovine CD14 stimulates growth (induce DNA  
 CC synthesis) in resting murine spleen cells and is 200 times more  
 CC effective than lipopolysaccharide (LPS), with the effect unaffected by  
 CC presence of serum. It also induces Ig secretion and a partial isotype  
 CC switch from IgM to IgG, in absence of T cells.  
 SQ Sequence 366 AA:

Query Match 84.6%; Score 44; DB 1; Length 366;  
 Best Local Similarity 90.9%; Pred. No. 0.26;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KSLSLRKRYR 11  
 Db 83 KSLSLRKRYR 93

RESULT 4  
 ID W89921 standard; Protein: 500 AA.  
 AC W89921;  
 DT 18-FEB-1999 (first entry)  
 DE Antigen 6 from cluster 38f.  
 KM Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;  
 KM peptic ulcer; gastric adenocarcinoma; gastric lymphoma.  
 OS Helicobacter pylori.  
 PN W09849314-A2.  
 PD 05-NOV-1998.  
 PE 27-APR-1998; U08487.  
 PR 14-OCT-1997; US-061958.  
 PR 25-APR-1997; US-045107.  
 PA (GENE-) GENELABS TECHNOLOGIES INC.  
 PI Chow TP, Fry KE, Lim MY, McAttee CP;  
 DR MPI: 99-009433/01.  
 PT New Helicobacter pylori antigens and related nucleic acid sequences  
 PT - useful in serological diagnosis and protective vaccines, providing  
 PT long-lasting immune response  
 PS Claim 1; Page 250-251; 402pp; English.  
 CC The present sequence represents a Helicobacter pylori antigenic protein  
 CC that is characterised by immunoreactivity with H. pylori-positive  
 CC antisera. The proteins are highly immunogenic and induce a long-lasting  
 CC immune response that persists even after antimicrobial treatment. In  
 CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are  
 CC highly sensitive and specific. The specification also describes 69  
 CC previously unrecognised immunogenic cluster families. H. pylori antigens  
 CC are used to detect H. pylori-specific antibodies, for diagnosing  
 CC infection or to confirm eradication of infection, and in vaccines to  
 CC protect against H. pylori infection and related diseases (gastritis,  
 CC peptic ulcer, gastric adenocarcinoma/lymphoma).  
 SQ Sequence 500 AA:

Query Match 69.2%; Score 36; DB 1; Length 500;  
 Best Local Similarity 63.6%; Pred. No. 14;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KSLSLRKRYR 11  
 Db 42 KSLSLRKRYR 52

RESULT 5  
 ID W89844 standard; Protein: 499 AA.  
 AC W89844;  
 DT 18-FEB-1999 (first entry)  
 DE Protein encoded by clone g9 ORF4.  
 KM Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;  
 KM peptic ulcer; gastric adenocarcinoma; gastric lymphoma.  
 OS Helicobacter pylori.  
 PN W09849314-A2.  
 PD 05-NOV-1998.  
 PE 27-APR-1998; U08487.  
 PR 14-OCT-1997; US-061958.  
 PR 25-APR-1997; US-045107.  
 PA (GENE-) GENELABS TECHNOLOGIES INC.  
 PI Chow TP, Fry KE, Lim MY, McAttee CP;  
 DR MPI: 99-009433/01.  
 DR N-PSDB; V90772.  
 PT New Helicobacter pylori antigens and related nucleic acid sequences  
 PT - useful in serological diagnosis and protective vaccines, providing  
 PT long-lasting immune response  
 PS Claim 15; Page 191-192; 402pp; English.  
 CC The present sequence represents a Helicobacter pylori antigenic protein  
 CC that is characterised by immunoreactivity with H. pylori-positive

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:49:51; Search time 135.78 Seconds  
(without alignments)

1.919 Million cell updates/sec

Title: US-09-124-280A-22

Perfect score: 52

Sequence: 1 KSLSKRLTYR 11

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	52	100.0	11	1	R71793	Peptide neutralist
2	52	100.0	11	1	W21610	Antibiotic potent
3	44	84.6	36	1	W60855	Mouse CD14 protein
4	36	69.2	500	1	W89921	Antigen 6 from clu
5	36	69.2	499	1	W89844	Protein encoded by
6	33	63.5	524	1	P90113	Rabies G protein G
7	33	63.5	524	1	P05077	Sali restriction p
8	32	61.5	1722	1	W38429	Human dendritic ce
9	31	59.6	420	1	W10168	Human regulator of
10	31	59.6	552	1	W50009	Murine hyaluronan
11	31	59.6	241	1	W67668	Human HOEFC11 prot
12	31	59.6	799	1	W86352	Human DNAX toll-11
13	31	59.6	837	1	W86361	Human DNAX toll-11
14	31	59.6	428	1	W88439	Arabidopsis pyruva
15	31	59.6	131	1	Y12477	Human 5' EST seque
16	30	57.7	1403	1	R10890	lXta::lacZ fusion
17	30	57.7	1022	1	P50097	Fusion protein of
18	30	57.7	1199	1	R26188	HYTA antigen. Inse
19	30	57.7	1805	1	R27204	Rat nestin. Diagno
20	30	57.7	1618	1	R27205	Human nestin. Diag
21	30	57.7	1121	1	R36728	Beat-galactosidase
22	30	57.7	1217	1	R49832	Rat nestin protein
23	30	57.7	1805	1	R60126	Human nestin prote
24	30	57.7	1618	1	R60127	Methyl tetrahydrop
25	30	57.7	456	1	W38570	Leukotoxin 101. Ch
26	30	57.7	1403	1	W79574	H. pylori GHPo 147
27	30	57.7	338	1	W98827	Protein encoded by
28	30	57.7	1015	1	W97241	Protein encoded by
29	30	57.7	74	1	Y11151	S. pneumoniae nitr
30	30	57.7	50	1	R05130	CF peptide encode
31	29	55.8	1822	1	R55273	Beta subunit of in
32	29	55.8	1822	1	R55273	Calcitonin gene-re
33	29	55.8	461	1	R32220	Human calcitonin r
34	29	55.8	472	1	R32753	Delta-endotoxin MI
35	29	55.8	287	1	R88002	Human DEC-205. Lig
36	29	55.8	1743	1	W00644	Clarika breweri S-
37	29	55.8	870	1	W17080	Platenolide syntha
38	29	55.8	1996	1	W23607	Platenolide syntha
39	29	55.8	1996	1	W23717	Selenomonas rumina
40	29	55.8	346	1	W26790	Bankia gouldi g1yc
41	29	55.8	956	1	W49874	Human secreted pro
42	29	55.8	335	1	W74727	linalool synthase
43	29	55.8	870	1	W73485	

## ALIGNMENTS

44	29	55.8	216	1	W98180	Arabidopsis cyclin
45	28	53.8	228	1	W98418	H. pylori GHPo 319

RESULT	1	2
ID	R71793	W21610
AC	R71793	W21610
DT	01-OCT-1995 (first entry)	26-AUG-1997 (first entry)
DE	Peptide neutralising toxicity of Lipid A.	Antibiotic potentiating peptide #22.
KW	Endotoxin; Lipid A; septic shock; lipopolysaccharide.	Potentiate; antibiotic; microbial infection; lipopolysaccharide;
OS	Synthetic.	KW permeability; outer bacterial membrane.
PN	W09503527-A.	OS Synthetic.
PD	02-FEB-1995.	PN W09638165-A1.
PF	21-JUL-1994; E02413.	PD 05-DEC-1996.
PR	26-JUL-1993; US-097830.	PF 29-MAY-1996; E02313.
PA	(BIOS-) BIOSYNTH SRL.	PR 31-MAY-1995; US-456112.
PI	PORRO M.	PA (BIOS-) BIOSYNTH SRL.
DR	WPI; 95-075190/10.	PI PORRO M.
PT	New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids	DR WPI; 97-034095/03.
PS	Claim 24: Page 22: 26pp; English.	PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic requy
CC	New peptides are claimed which are linear or cyclic peptides which include units of formula: (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.	PS Claim 26: Page 26: 37pp; English.
CC	The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable nutrient solutions, etc.	CC The sequences given in W21589-633 represent peptides which act to potentiate the activity of an antibiotic when they are co-administered with the antibiotic. Compositions containing these peptides are used to treat or prevent microbial infections. These peptides bind to
CC	The present sequence is a specifically claimed example of the new peptides.	CC to treat or prevent microbial infections. These peptides bind to
SC	Sequence 11 AA:	

QY	1 KSLSKRLTYR 11	Query Match 100.0%; Score 52; DB 1; Length 11;
DB	1 KSLSKRLTYR 11	Best Local Similarity 100.0%; Pred. No. 0.00016;
		Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT	2
ID	W21610
AC	W21610
DT	26-AUG-1997 (first entry)
DE	Antibiotic potentiating peptide #22.
KW	Potentiate; antibiotic; microbial infection; lipopolysaccharide;
OS	permeability; outer bacterial membrane.
PN	Synthetic.
PD	W09638165-A1.
PF	05-DEC-1996.
PR	29-MAY-1996; E02313.
PA	31-MAY-1995; US-456112.
PI	(BIOS-) BIOSYNTH SRL.
DR	PORRO M.
PT	WPI; 97-034095/03.
PS	Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic requy
CC	Claim 26: Page 26: 37pp; English.
CC	The sequences given in W21589-633 represent peptides which act to potentiate the activity of an antibiotic when they are co-administered with the antibiotic. Compositions containing these peptides are used to treat or prevent microbial infections. These peptides bind to
CC	to treat or prevent microbial infections. These peptides bind to



**THIS PAGE BLANK (USPTO)**

Search completed: September 7, 1999, 22:47:26  
Job time: 7946 sec

```

FT TRANSMEM 461 481 POTENTIAL.
FT TRANSMEM 514 534 POTENTIAL.
FT TRANSMEM 549 569 POTENTIAL.
FT TRANSMEM 581 601 POTENTIAL.
FT TRANSMEM 606 626 POTENTIAL.
FT TRANSMEM 642 662 POTENTIAL.
FT TRANSMEM 723 743 POTENTIAL.
FT TRANSMEM 946 966 POTENTIAL.
FT TRANSMEM 980 1000 POTENTIAL.
FT TRANSMEM 1021 1041 POTENTIAL.
FT TRANSMEM 1075 1095 POTENTIAL.
FT TRANSMEM 1148 1168 POTENTIAL.
FT TRANSMEM 1193 1213 POTENTIAL.
FT TRANSMEM 1274 1294 POTENTIAL.
FT TRANSMEM 1302 1322 POTENTIAL.
FT TRANSMEM 1331 1351 POTENTIAL.
FT TRANSMEM 1358 1378 POTENTIAL.
FT TRANSMEM 1393 1413 POTENTIAL.
FT TRANSMEM 1486 1506 POTENTIAL.
FT TRANSMEM 1631 1651 POTENTIAL.
SO SEQUENCE 1854 AA; 213163 MM; A2F098BA CRC32;

```

```

Query Match      72.5%; Score 29; DB 3; Length 1854;
Best Local Similarity 87.5%; Pred. No. 6; 7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 KALVRRL 9
Db 1049 KALVRRL 1056

```

```
RESULT 11
```

```
075099 PRELIMINARY; PRT: 479 AA.
```

```

AC 075099;
D8 01-NOV-1998 (Tremblrel. 08, Created)
D7 01-NOV-1998 (Tremblrel. 08, Last sequence update)
D7 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE CHONDROITIN 6-SULFOTRANSFERASE (EC 2.8.2.17) (C6ST).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
RC TISSUE=FETAL BRAIN;
RX MEDLINE: 98390270.
RA FUKUTA M., KOBAYASHI Y., UCHIMURA K., KIMATA K., HABUCHI O.;
RT "Molecular cloning and expression of human chondroitin
RL Blochm. Blophys. Acta 1399:57-61(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA TSUTSUMI K., SHIMAWA H., KITAGAWA H., SUGAHARA K.;
RT "Functional expression and genomic structure of human chondroitin
RL 6-sulfotransferase."

```

```

CC CC -1- FUNCTION: CATALYZES THE SULFATION OF CHONDROITIN AND KERATAN
CC CC SULFATE.
CC CC -1- CATALYTIC ACTIVITY: 3'-PHOSPHOADENYLYLSULFATE + CHONDROITIN =
CC CC ADENOSINE 3',5'-BISPHOSPHATE + CHONDROITIN 6'-SULFATE.
CC CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI MEMBRANE
CC CC (BY SIMILARITY).
CC CC -1- TISSUE SPECIFICITY: FOUND AT HIGHEST LEVELS IN SKELETAL MUSCLE.
CC CC ALSO PRESENT AT LOWER LEVELS IN HEART, PLACENTA AND PANCREAS.

```

```

CC CC EMBL: AB012192; BAA32576.1; -.
CC CC EMBL: AB017915; BAA36348.1; -.
CC CC TRANSFERASE; Transmembrane; Signal-anchor; Glycoprotein; Golgi stack.

```

```

FT TRANSMEM 24 38 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 39 458 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN (BY SIMILARITY).
FT LUMENAL, CATALYTIC (POTENTIAL).

```

```

FT CARBOHYD 63 63 POTENTIAL.
FT CARBOHYD 74 74 POTENTIAL.
FT CARBOHYD 96 96 POTENTIAL.
FT CARBOHYD 256 256 POTENTIAL.
FT CARBOHYD 420 420 POTENTIAL.
FT CARBOHYD 464 464 POTENTIAL.
FT CARBOHYD 387 387 POTENTIAL.
FT VARIANT 443 443 P -> R (IN PLACENTA).
FT VARIANT 443 443 P -> A (IN PLACENTA).
SO SEQUENCE 479 AA; 54672 MM; FB77CAD0 CRC32;

```

```

Query Match      72.5%; Score 29; DB 4; Length 479;
Best Local Similarity 55.6%; Pred. No. 1; 9e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 VKALVRRL 9
Db 272 LKAVIRRL 280

```

```
RESULT 12
```

```
088199 PRELIMINARY; PRT: 472 AA.
```

```

AC 088199;
D7 01-NOV-1998 (Tremblrel. 08, Created)
D7 01-NOV-1998 (Tremblrel. 08, Last sequence update)
D7 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE CHONDROITIN 6-SULFOTRANSFERASE.
GN C6ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; AND 129; TISSUE=SPLEEN;
RX MEDLINE: 98259835.
RA UCHIMURA K., KADOMATSU K., FAN Q.-W., MURAMATSU H., KUROSAWA N.,
RA KANAME T., YAMAMURA K., FUKUTA M., HABUCHI O., MURAMATSU T.;
RT "Mouse chondroitin 6-sulfotransferase: molecular cloning,
RL characterization and chromosomal mapping."

```

```

DR EMBL: AB008937; BAA29054.1; -.
DR EMBL: AB008938; BAA29055.1; -.
KM TRANSFERASE.
SO SEQUENCE 472 AA; 53997 MM; 1CFCC83B CRC32;

```

```

Query Match      72.5%; Score 29; DB 11; Length 472;
Best Local Similarity 55.6%; Pred. No. 1; 8e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 VKALVRRL 9
Db 266 LKAVIRRL 274

```

```
RESULT 13
```

```
077670 PRELIMINARY; PRT: 37 AA.
```

```

AC 077670;
D7 01-NOV-1996 (Tremblrel. 01, Created)
D7 01-NOV-1996 (Tremblrel. 01, Last sequence update)
D7 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93219569.
RA GOMEZ CARRILLO M., PICCARDO C., LIBONATTI O.;
RT "Molecular analysis of the principal neutralization epitope (V3
loop) of human immunodeficiency virus type 1 in Argentina."

```

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA SENGAR A.S., WANG W., BISHAY J., COHEN S., EGAN S.E.;  
 RT "The EH and SH domain Ese proteins regulate endocytosis by linking  
 to dynamin and Eps15."  
 RL EMBL J. 18:1159-1171(1999).  
 DR EMBL; AF132481; AAD19749.1;  
 SQ SEQUENCE 1714 AA; 194284 MW; 37D424DF CRC32;

Query Match 75.0%; Score 30; DB 11; Length 1714;  
 Best Local Similarity 75.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRR 8  
 Db 1289 LKALRVRR 1296

RESULT 7  
 ID 056284 PRELIMINARY; PRT; 138 AA.  
 AC 056284;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)  
 DE ORF-2.  
 OS Thibacillus ferrooxidans.  
 OC Bacteria; Proteobacteria; gamma subdivision.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-E-15;  
 RX MEDLINE; 92140035.  
 RA INQUE C., SUGAWARA K., KUSANO T.;  
 RT "The merR regulatory gene in Thibacillus ferrooxidans is spaced  
 apart from the mer structural genes."  
 RL MOJ. Microbiol. 5:2707-2718(1991).  
 DR EMBL; X57326; CAA40600.1;  
 SQ SEQUENCE 138 AA; 15485 MW; C120EDB9 CRC32;

Query Match 72.5%; Score 29; DB 2; Length 138;  
 Best Local Similarity 62.5%; Pred. No. 58;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRR 8  
 Db 125 VOAMRIRR 132

RESULT 8  
 ID 060027 PRELIMINARY; PRT; 544 AA.  
 AC 060027;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)  
 OC 01-NOV-1998 (TREMblrel. 08, Last annotation update)  
 DE ALPHA-GLUCOSIDASE (EC 3.2.1.20) (MALTASE) (GLUCOINVERTASE)  
 DE (GLUCOSIDOSUCRASE) (MALTASE-GLUCAMYLASE)  
 DE (LYSOSOMAL ALPHA-GLUCOSIDASE) (ACID MALTASE).  
 GN AGLA.  
 OS Thermomonospora curvata.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptosporangineae; Thermomonosporaceae;  
 OC Thermomonospora.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CM3352;  
 RA JANDA L., TICHY P., PETRICEK M.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING  
 CC 1,4-LINKED D-GLUCOSE RESIDUES WITH RELEASE OF D-GLUCOSE.  
 DR EMBL; U17917; AAA57313.1; -

DR PFAM; PF00128; alpha-amylase; 1.  
 KW Hydrolase; Glycosidase.  
 SQ SEQUENCE 544 AA; 60874 MW; 51658607 CRC32;

Query Match 72.5%; Score 29; DB 2; Length 544;  
 Best Local Similarity 85.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 KALRVRR 8  
 Db 465 KALRVRR 471

RESULT 9  
 ID 0929H1 PRELIMINARY; PRT; 581 AA.  
 AC 0929H1;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)  
 DE SIMILAR TO YJCE-ECOLI.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PAOI;  
 RA NASHIMOTO H.;  
 RT "P.aeruginosa PAOI."  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB024601; BAA75912.1;  
 SQ SEQUENCE 581 AA; 63624 MW; 09AA4939 CRC32;

Query Match 72.5%; Score 29; DB 2; Length 581;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VKALRVRR 9  
 Db 535 LKALRVRR 543

RESULT 10  
 ID 014234 PRELIMINARY; PRT; 1854 AA.  
 AC 014234;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)  
 DE PROBABLE SODIUM CHANNEL PROTEIN C6P6.01.  
 GN SPAC6P6.01.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Arhizosmycetaceae;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA GENTLES S., CHURCHER C.M., BARRELL B.G., RAINDREAN M.A., WOOD V.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: TO MAMMALIAN SODIUM AND CALCIUM CHANNELS.  
 DR EMBL; Z96981; CAB11726.1; -  
 DR PFAM; PF00520; Ion.trans. 3.  
 KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel.  
 FT DOMAIN 2  
 FT TRANSMEM 184 204  
 FT TRANSMEM 220 240  
 FT TRANSMEM 274 294  
 FT TRANSMEM 328 348  
 FT TRANSMEM 427 447  
 FT TRANSMEM POTENTIAL.  
 FT TRANSMEM POTENTIAL.

RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,  
 RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,  
 RA HALSTORF K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.,  
 RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,  
 RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIEFEN L., RILES L.,  
 RA TACH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,  
 RA WILSON R., WATERSTON R.,  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972):  
 RA WATERSTON R.,  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 2353-2958 FROM N.A.  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA BENES V., RECHMANN S., NENTWICH U., SCHWAGER C., ANSORGE W., VOSS H.,  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA MIPS:  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U53880; AAB67591.1; -.  
 DR EMBL: 273259; CAA97647.1; -.  
 DR SCD: L0000557; GAIL.  
 SO SEQUENCE 2958 AA; 338254 MW; 60D78E22 CRC32;

Query Match 75.0%; Score 30; DB 3; Length 2958;  
 Best Local Similarity 87.5%; Pred. No. 6.4e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKALVR 8  
 DB 2857 VKALVR 2864  
 ID 095062 PRELIMINARY; PRT; 464 AA.  
 AC 095062;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE SHP18-LIKE WASP ASSOCIATED PROTEIN (FRAGMENT).  
 GN SWAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA WILSON L.A., FIELDS D., CRUZ L., FRIESEN J., SIMINOVITCH K.A.,  
 RT "SHP18-like wasp associated protein (SWAP): A multiple SH3 domain  
 containing protein that interacts with WASP."  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF001630; AAD00899.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 464  
 SO SEQUENCE 464 AA; 51383 MW; 257CE57D CRC32;

Query Match 75.0%; Score 30; DB 4; Length 464;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALVR 8  
 DB 447 LKALVR 454  
 ID 095062 PRELIMINARY; PRT; 464 AA.  
 AC 095062;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE SHP18-LIKE WASP ASSOCIATED PROTEIN (FRAGMENT).  
 GN SWAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA WILSON L.A., FIELDS D., CRUZ L., FRIESEN J., SIMINOVITCH K.A.,  
 RT "SHP18-like wasp associated protein (SWAP): A multiple SH3 domain  
 containing protein that interacts with WASP."  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF001630; AAD00899.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 464  
 SO SEQUENCE 464 AA; 51383 MW; 257CE57D CRC32;

RESULT 4  
 095216 PRELIMINARY; PRT; 1721 AA.  
 AC 095216;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE INTERSECTIN LONG FORM.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN.  
 RX MEDLINE: 99017974.  
 RA GUTFRONT M., SCOTT H.S., CHEN H., SCHEBESTA A., ROSSIER C.,  
 RA ANTONARAKIS S.E.,  
 RT "Two isoforms of a human intersectin (ITSN) protein are produced by  
 brain-specific alternative splicing in a stop codon."  
 RL Genomics 53:369-376(1998).  
 DR EMBL: AF064244; AAC78611.1; -.  
 SO SEQUENCE 1721 AA; 195559 MW; A5A69515 CRC32;

Query Match 75.0%; Score 30; DB 4; Length 1721;  
 Best Local Similarity 75.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALVR 8  
 DB 1296 LKALVR 1303  
 ID 097015 PRELIMINARY; PRT; 257 AA.  
 AC 097015;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE L7610.13 PROTEIN.  
 GN L7610.13.  
 OS Leishmania major.  
 OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmania.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA OLIVER K., MURPHY L., QUAIL M., LAWSON D., HARRIS D., RAJANDREAM M.,  
 RA IYENS A., BARRELL B.,  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL034356; CAA22248.1; -.  
 SO SEQUENCE 257 AA; 27995 MW; A2381389 CRC32;

Query Match 75.0%; Score 30; DB 5; Length 257;  
 Best Local Similarity 66.7%; Pred. No. 63;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VKALVR 9  
 DB 82 VRLARVR 90  
 ID 0920R4 PRELIMINARY; PRT; 1714 AA.  
 AC 0920R4;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE ESRL PROTEIN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RESULT 6  
 0920R4 PRELIMINARY; PRT; 1714 AA.  
 AC 0920R4;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE ESRL PROTEIN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:35:00 ; Search time 160.2 Seconds  
(without alignments)  
3.458 Million cell updates/sec

Title: US-09-124-280A-21

Perfect score: 40  
Sequence: 1 VKALRVRL 9

Scoring table: BLOSUM62

Searched: 201082 segs, 61543640 residues

Database :

SPTREMBL.10:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	33	82.5	462	2	053116	053116 rhodococcus
2	30	75.0	2958	3	02150	02150 saccharomyc
3	30	75.0	464	4	095062	095062 homo sapien
4	30	75.0	1721	4	095216	095216 homo sapien
5	30	75.0	257	5	097015	097015 leishmania
6	30	75.0	1714	11	092084	092084 mus musculu
7	29	72.5	138	2	056284	056284 thlobacillu
8	29	72.5	544	2	060027	060027 thermomonos
9	29	72.5	581	2	0929H1	0929H1 pseudomonas
10	29	72.5	1854	3	014234	014234 schizosacch
11	29	72.5	479	4	075039	075039 homo sapien
12	29	72.5	472	11	088199	088199 mus musculu
13	29	72.5	37	12	07670	07670 human immun
14	29	72.5	286	13	091988	091988 xiphophorus
15	28	70.0	1004	1	028412	028412 archaeglob
16	28	70.0	310	1	028545	028545 archaeglob
17	28	70.0	814	1	052767	052767 shewanella
18	28	70.0	125	2	054409	054409 streptomyces
19	28	70.0	483	2	056832	056832 xanthomonas
20	28	70.0	392	2	030529	030529 pseudomonas
21	28	70.0	390	2	050279	050279 pseudomonas
22	28	70.0	398	2	051136	051136 borrellia bu
23	28	70.0	229	2	066552	066552 aquifex aeo
24	28	70.0	478	2	066610	066610 aquifex aeo
25	28	70.0	295	2	095534	095534 pseudomonas
26	28	70.0	378	2	083825	083825 treponema p
27	28	70.0	200	2	086646	086646 streptomyces
28	28	70.0	134	2	087054	087054 vibrio chol
29	28	70.0	1435	3	003291	003291 saccharomyc

30	28	70.0	684	4	P78424	P78424 homo sapien
31	28	70.0	648	4	P78425	P78425 homo sapien
32	28	70.0	279	5	O23486	O23486 caenorhabdi
33	28	70.0	2257	10	O40326	O40326 medicago sa
34	28	70.0	512	11	O61745	O61745 mus musculu
35	28	70.0	512	11	O61364	O61364 mus musculu
36	28	70.0	506	11	O62662	O62662 ratius norv
37	28	70.0	237	12	O66623	O66623 equine herp
38	28	70.0	507	12	O9YZA0	O9YZA0 rinderpest
39	28	67.5	776	1	O27582	O27582 methanobact
40	27	67.5	177	1	O27666	O27666 methanobact
41	27	67.5	324	1	O28924	O28924 archaeglob
42	27	67.5	141	2	O30539	O30539 pseudomonas
43	27	67.5	3588	2	O66070	O66070 bacillus li
44	27	67.5	3589	2	O69246	O69246 bacillus li
45	27	67.5	4114	11	O54796	O54796 mus musculu

## ALIGNMENTS

RESULT 1  
ID 053116 PRELIMINARY: PRT; 462 AA.  
AC 053116;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)  
DE ENAMIDOMERASE-SELECTIVE AMIDASE.  
GN AMDA.  
OS Rhodococcus sp.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MAYAUX J.F., GERBELAUD E., SOUBRIER F., YEH P., BLANCHE F., PETRE D.,  
RA YEH P.;  
RL Submitted (OCT-1991) to the EMBL/GenBank/DBJ databases.  
DR EMBL: M74531; AAA26183.1; -  
DR EMBL: A19131; CAA01439.1; -  
DR PRAM: PF01425; Amidase; 1.  
SQ SEQUENCE 462 AA; 48609 MW; 6E8FF741 CRC32;

Query Match 82.5%; Score 33; DB 2; Length 462;  
Best Local Similarity 77.8%; Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
DB 347 IKALRVRL 355

RESULT 2  
ID 012150 PRELIMINARY: PRT; 2958 AA.  
AC 012150;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DT 01-NOV-1996 (TREMblrel. 01, Last annotation update)  
DE CHROMOSOME XII COSMID 9449.  
GN GAAL OR L9449.14.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomyces.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C (AB972);  
RA GEISEL C.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RC SEQUENCE FROM N.A.  
RC STRAIN-S288C (AB972);

Search completed: September 7, 1999, 23:58:59  
 Job time: 511 sec

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U14003; G537164; -  
 DR EMBL: AE000503; G1790779; -  
 DR EMBL: D13329; D1003095; -  
 DR ECGENE: EG30248; UXUB  
 DR PROSITE: PS00974; MANNITOL\_DHENASE; 1.  
 DR PFAM: PF01232; Mannitol\_dh; 1.  
 KW OXIDOREDUCTASE; NAD.  
 FT NP\_BIND 25 36 NAD (BY SIMILARITY).  
 SO SEQUENCE 486 AA; 53580 MW; 2CC0EF76 CRC32;

Query Match 72.5%; Score 29; DB 1; Length 486;  
 Best Local Similarity 75.0%; Pred. No. 40;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRR 8  
 1:|||||  
 DB 166 VEALRLRR 173

RESULT 15

VG59\_HSV11 STANDARD; PRT; 345 AA.  
 AC 000138;

DT 01-DEC-1992 (REL. 24, CREATED)

DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)

DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL GENE 59 MEMBRANE PROTEIN.

OS ICTALURID HERPESVIRUS 1 (CHANNEL CATFISH VIRUS) (CCV).  
 OC VIRUSES: DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;

OC UNCLASSIFIED HERPESVIRIDAE.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-AUBURN 1;

RX MEDLINE: 92087490.  
 RA DAVISON A.J.;

RT "Channel catfish virus: a new type of herpesvirus."  
 RL VIROLOGY 186:9-14(1992).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: M75136; G331269; -

DR PIR: F36792; MMBE17.  
 KM HYPOTHETICAL PROTEIN; TRANSMEMBRANE.

FT TRANSMEM 46 63 POTENTIAL.  
 FT TRANSMEM 101 118 POTENTIAL.

FT TRANSMEM 147 165 POTENTIAL.  
 FT TRANSMEM 265 286 POTENTIAL.

SO SEQUENCE 345 AA; 38135 MW; CAF75BD5 CRC32;

Query Match 72.5%; Score 29; DB 1; Length 345;  
 Best Local Similarity 66.7%; Pred. No. 27;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VKALRVRR 9  
 1:|||||  
 DB 286 IRALRRRL 294

```

CC CAROPHYLLIDAE: NEPENTHINEAE; DROSERACEAE; DROSERA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92397008.
RA ALBERT V.A., WILLIAMS S.E., CHASE M.W.;
RT "Carnivorous plants: phylogeny and structural evolution.";
RL SCIENCE 257:1491-1495(1992).
CC -1- FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
CC 2 3-PHOSPHO-D-GLYCERATE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L01914; G424049; -
CC PROSITE: PS00157; RUBISCO_LARGE; 1.
CC PFM: PF00016; RUBISCO_LARGE; 1.
CC MENDEL: 2944; DROE; rbcL; 1.
CC HSSP: P00875; LAUS.
CC PHOTOSYNTHESIS: CARBON DIOXIDE FIXATION; PHOTORESPIRATION;
CC LYASE: OXIDOREDUCTASE; MONOOXYGENASE; CHLOROPLAST.
CC NON_TER 1
CC ACT_SITE 1
CC SEQUENCE 466 AA; 51813 MW; 915F38E1 CRC32;
SQ

```

Query Match 72.5%; Score 29; DB 1; Length 466;  
 Best Local Similarity 87.5%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KALRYRL 9  
 11111111  
 Db 119 KALRYRL 126

RESULT 13  
 RBL\_VERBO STANDARD; PRT; 443 AA.  
 AC P36490;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO  
 DE LARGE SUBUNIT) (FRAGMENT).  
 GN RBL.  
 OS VERBENA BONARIENSIS.  
 OC CHLOROPLAST.  
 OC EUKARYOTA: VIRIDIPANTAE: STREPTOPHYTA: EMBRYOPHYTA: TRACHEOPHYTA;  
 CC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;  
 CC ASTERIDAE: GENTIANANAE; LAMIALES; VERBENACEAE; VERBENA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA OLMSTAD R.G., BREMER B., SCOTT K.M., PALMER J.D.;  
 RT "A parsimony analysis of the Asteridae sensu lato based on rbcL  
 RT sequences";  
 RL ANN. MO. BOT. GARD. 80:700-722(1993).  
 CC -1- FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF  
 CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC

```

CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
CC 2 3-PHOSPHO-D-GLYCERATE.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L14412; G349157; -
CC PROSITE: PS00157; RUBISCO_LARGE; 1.
CC PFM: PF00016; RUBISCO_LARGE; 1.
CC MENDEL: 2882; VERBO; rbcL; 1.
CC HSSP: P00876; 3RUB.
CC PHOTOSYNTHESIS: CARBON DIOXIDE FIXATION; PHOTORESPIRATION;
CC LYASE: OXIDOREDUCTASE; MONOOXYGENASE; CHLOROPLAST.
CC NON_TER 1
CC ACT_SITE 1
CC SEQUENCE 443 AA; 49136 MW; 3C6F1CAB CRC32;
SQ

```

Query Match 72.5%; Score 29; DB 1; Length 443;  
 Best Local Similarity 87.5%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KALRYRL 9  
 11111111  
 Db 94 KALRYRL 101

RESULT 14  
 UXUB\_ECOLI STANDARD; PRT; 486 AA.  
 AC P39160;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE D-MANNONATE OXIDOREDUCTASE (EC 1.1.1.57) (FRUCTURONATE REDUCTASE).  
 GN UXUB.  
 OS ESCHERICHIA COLI.  
 OC BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION: ENTEROBACTERIACEAE;  
 CC ESCHERICHIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-K12 / MG1655;  
 RX MEDLINE: 95334362.  
 RA BURLAND V.D., PLUNKETT G. III, SOFIA H.J., DANIELS D.L.,  
 RA BRATNER F.R.;  
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the  
 RT region from 92.8 through 100 minutes";  
 RL NUCLEIC ACIDS RES. 23:2105-2119(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-K12 / W3110.  
 RA MIZOBUCHI K.;  
 RL SUBMITTED (SEP-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- CATALYTIC ACTIVITY: D-MANNONATE + NAD(+) = D-FRUCTURONATE + NADH.  
 CC -1- SIMILARITY: BELONGS TO THE MANNITOL DEHYDROGENASES FAMILY.  
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its



FT NP\_BIND 107 114 GTP (BY SIMILARITY).  
 FT NP\_BIND 188 192 GTP (BY SIMILARITY).  
 FT NP\_BIND 247 250 GTP (BY SIMILARITY).  
 SO SEQUENCE 451 AA: 50124 MW: 028734AD CRC32:

Query Match 75.0%: Score 30; DB 1; Length 451;  
 Best Local Similarity 55.6%: Pred. No. 22;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VKALVRRL 9  
 :|||:|:  
 Db 391 IKASRIRRI 399

RESULT 10  
 DNBI\_EBV  
 ID DNBI\_EBV STANDARD: PRT; 1128 AA.  
 AC P03227;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE MAJOR DNA-BINDING PROTEIN.  
 GN BALF2.  
 OS EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).  
 OC VIRUSES: DSDNA VIRUSES, NO RNA STAGE: HERPESVIRIDAE;  
 CC GAMMAHERPESVIRINAE; LYMPHOCRYPTOVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 84270667.  
 RA BAKER R., BARKIER A.T., BIGGIN M.D., DEININGER P.L., FARRELL P.J.,  
 RA GIBSON T.J., HATEFUL G., HUDSON G.S., SATCHELL S.C., SEGUIN C.,  
 RA TUFENNEL P.S., BARRELL B.G.:  
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome."  
 RT NATURE 310:207-211(1984).  
 RL NATURE 310:207-211(1984).  
 CC -1- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA  
 CC REPLICATION.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES' DNA-BINDING PROTEIN.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: V01555; G59157; -.  
 DR PIR: A03791; Q0BE47.  
 DR PIR: S33057; S33057.  
 DR PIR: P00747; V1R1\_DNA\_bp; 1.  
 DR DNA-BINDING: DNA REPLICATION; ZINC-FINGER; NUCLEAR PROTEIN;  
 KW EARLY PROTEIN.  
 SO SEQUENCE 1128 AA: 123121 MW: F2C73149 CRC32:

Query Match 72.5%: Score 29; DB 1; Length 1128;  
 Best Local Similarity 75.0%: Pred. No. 1e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 KALVRRL 9  
 :|||:|:  
 Db 223 QALVRRL 230

RESULT 11  
 PHSI\_DESBA  
 ID PHSI\_DESBA STANDARD: PRT; 513 AA.  
 AC P13065;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE PERIPLASMIC [NIFESF] HYDROGENASE LARGE SUBUNIT (EC 1.18.99.1) (NIFESF  
 DE HYDROGENLYASE LARGE CHAIN).  
 OS DESULFOVIBRIO BACILLATUS (DESULFOMICROBIUM BACILLATUS).  
 OC BACTERIA; PROTEOBACTERIA; DELTA SUBDIVISION; DESULFOMICROBIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 88058744.  
 RA MENON N.K., PECK H.D. JR., LE GALL J., PRZYBYLA A.E.:  
 RT "Cloning and sequencing of the genes encoding the large and small  
 RT subunits of the periplasmic (Nifese) hydrogenase of Desulfovibrio  
 RT bacillatus."  
 RT J. BACTERIOL. 169:5401-5407(1987).  
 RL [2]

RP REVISIONS.  
 RA MENON N.K., PECK H.D. JR., LE GALL J., PRZYBYLA A.E.:  
 RL J. BACTERIOL. 170:4429-4429(1988).  
 CC -1- CATALYTIC ACTIVITY: 2 REDUCED FERREDOXIN + 2 H(+) = 2 OXIDIZED  
 CC FERREDOXIN + H(2).  
 CC -1- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CENTERS AND ONE  
 CC SELENOCYSTEINE.  
 CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.  
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.  
 CC -1- PERHAPS THE LEADER OF THE SMALL SUBUNIT SERVES AS A TRANSPORT  
 CC VEHICLE FOR BOTH SUBUNITS.  
 CC -1- SIMILARITY: BELONGS TO THE [NIFESF]/[NIFESF] HYDROGENASE LARGE  
 CC SUBUNIT FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M18271; G145102; -.  
 DR PIR: A33101; H0DVLB.  
 DR PROSITE: P500507; NI\_HGENASE\_L\_1; 1.  
 DR PROSITE: P500508; NI\_HGENASE\_L\_2; 1.  
 DR PIR: P00374; Nifese\_Hases; 1.  
 DR HSP: P12944; 1FRV.  
 KW OXIDOREDUCTASE; PERIPLASMIC; NICKEL; SELENIUM; SELENOCYSTEINE.  
 FT INIT\_MET 0 0  
 FT METAL 70 70 NICKEL (POTENTIAL).  
 FT METAL 73 73 NICKEL (POTENTIAL).  
 FT BINDING 492 492 SELENIUM.  
 FT METAL 495 495 NICKEL (POTENTIAL).  
 SO SEQUENCE 513 AA: 56683 MW: 78233262 CRC32:

Query Match 72.5%: Score 29; DB 1; Length 513;  
 Best Local Similarity 55.6%: Pred. No. 42;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VKALVRRL 9  
 :|||:|:  
 Db 166 LKALEIRRI 174

RESULT 12  
 RBL\_DRORE  
 ID RBL\_DRORE STANDARD: PRT; 466 AA.  
 AC P28411;  
 DT 01-DEC-1992 (REL. 24, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO  
 DE LARGE SUBUNIT) (FRAGMENT).  
 GN RBL.  
 OS DROSERA REGIA.  
 OG CHLOROPLAST.  
 OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTES; EMBRYOPHYTES; TRACHEOPHYTES;  
 OC EUPHYLOPHYTES; SPERMATOPHYTES; MAGNOLIOPHYTES; EUDICOTYLEDONS;

```

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
GN TRAL. PROTEIN.
OS ESCHERICHIA COLI.
OG PLASMID INCP-BETA RP4.
OC BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION: ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-4.
RC STRAIN-HB101.
RX MEDLINE: 92190548.
RA ZIEGLIN G., PANSEGRAU W., STRACK B., BALZER D., KROEGER M.,
RA KRUET V., LANKA E.;
RT "Nucleotide sequence and organization of genes flanking the transfer
RT origin of promiscuous plasmid RP4.";
RL DNA SEQ. 1:303-327(1991).
CC -1- SIMILARITY: TO PLASMID R751 TRAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X54459; G42787; -
DR PIR: S23004; S23004.
KW PLASMID; CONJUGATION.
SO SEQUENCE 241 AA; 26566 MW; 364A829B CRC32;

Query Match 77.5%; Score 31; DB 1; Length 241;
Best Local Similarity 87.5%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KALVRRL 9
Db 50 KALVRRL 57

RESULT 8
TRLS ECOLI STANDARD; PRT; 241 AA.
ID 000187; P71197;
AC 01-FEB-1995 (REL. 31, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE TRAL. PROTEIN.
GN TRAL.
OS ESCHERICHIA COLI.
OC PLASMID INCP-BETA R751.
OC BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION: ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC THOMAS C.M.;
RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 1-48 FROM N.A.
RC STRAIN-HB101.
RX MEDLINE: 92190548.
RA ZIEGLIN G., PANSEGRAU W., STRACK B., BALZER D., KROEGER M.,
RA KRUET V., LANKA E.;
RT "Nucleotide sequence and organization of genes flanking the transfer
RT origin of promiscuous plasmid RP4.";
RL DNA SEQ. 1:303-327(1991).
CC -1- SIMILARITY: TO PLASMID INCP-BETA RP4 TRAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67194; G1572585; -
DR EMBL: X54458; G42663; -
DR PIR: S22998; S22998.
KW PLASMID; CONJUGATION.
SO SEQUENCE 241 AA; 26341 MW; FEA7633D CRC32;

Query Match 77.5%; Score 31; DB 1; Length 241;
Best Local Similarity 87.5%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KALVRRL 9
Db 50 KALVRRL 57

RESULT 9
SR54_METUA STANDARD; PRT; 451 AA.
ID SR54_METUA
AC 057565;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE PROBABLE SIGNAL RECOGNITION 54 KD PROTEIN (SRP54).
GN M0101.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA: EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1/DSM 2661 / ATCC 43067;
RX MEDLINE: 96337999.
RA BOLT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOBE J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUEHRMAN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLEIN H.-P., FRASER C.M., SMITH H.O., MOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL SCIENCE 273:1058-1073(1996).
CC -1- FUNCTION: BINDS TO THE SIGNAL SEQUENCE OF PRESECRETORY PROTEIN
CC WHEN THEY EMERGE FROM THE RIBOSOMES AND TRANSFERS THEM TO TRAM
CC (TRANSLOCATING CHAIN-ASSOCIATING MEMBRANE PROTEIN)
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- DOMAIN: HAS A TWO DOMAIN STRUCTURE. THE G-DOMAIN BINDS GTP. THE
CC M-DOMAIN BINDS THE 7S RNA IN PRESENCE OF SRP19 AND ALSO BINDS THE
CC SIGNAL SEQUENCE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67467; G1590878; -
DR TIGR: M0101; -
DR PROSITE: PS00300; SRP54; 1.
DR PFAM: PF00448; SRP54; 1.
DR HSPR: O07347; 1PFH
KW SIGNAL RECOGNITION PARTICLE; GTP-BINDING; RNA-BINDING.
FT DOMAIN 1 294 G-DOMAIN (BY SIMILARITY).
FT DOMAIN 295 451 M-DOMAIN (BY SIMILARITY).

```

RX MEDLINE; 93233655.  
 RA JIANG Y.P., WANG H., D'EUSTACHIO P., MUSACCHIO J.M., SCHLESSINGER J.,  
 RA SAP J.:  
 RT "Cloning and characterization of R-PTP-kappa, a new member of the  
 RT receptor protein tyrosine phosphatase family with a proteolytically  
 RT cleaved cellular adhesion molecule-like extracellular region."  
 RL MOL. CELL. BIOL. 13:2942-2951(1993).  
 CC -1- FUNCTION: REGULATION OF PROCESSES INVOLVING CELL CONTACT AND  
 CC ADHESION SUCH AS GROWTH CONTROL, TUMOR INVASION, AND METASTASIS.  
 CC FORMS COMPLEXES WITH BETA-CATENIN AND GAMMA-CATENIN/PLAKOGLOBIN.  
 CC BETA-CATENIN MAY BE A SUBSTRATE FOR THE CATALYTIC ACTIVITY OF PTP-  
 CC KAPPA.  
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN LIVER AND KIDNEY. LOWER LEVELS  
 CC IN LUNG, BRAIN AND HEART. NOT SEEN IN SPLEEN AND TESTIS.  
 CC -1- DEVELOPMENTAL STAGE: DEVELOPMENTALLY REGULATED WITH HIGHEST  
 CC EXPRESSION FOUND IN DEVELOPING AREAS OR IN AREAS CAPABLE OF  
 CC DEVELOPMENTAL PLASTICITY.  
 CC -1- PTM: THIS PROTEIN UNDERGOES PROTEOLYTIC PROCESSING.  
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IG-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L10106; G393772; -  
 DR PIR; A48066; A48066.  
 DR MGI; MGI:103310; PTPRK.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 2.  
 DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 2.  
 DR PROSITE; PS50055; TYR\_PHOSPHATASE\_PTP; 2.  
 DR PROSITE; PS00740; MAM\_1; 1.  
 DR PROSITE; PS50060; MAM\_2; 1.  
 DR PFAM; PF00041; fn3; 2.  
 DR PFAM; PF00047; 19; 1.  
 DR PFAM; PF00102; Y.phosphatase; 2.  
 DR PFAM; PF00629; MAM; 1.  
 DR HSP; P28827; IRM.  
 DR HYDROLASE; RECEPTOR; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
 KW IMMUNOGLOBULIN FOLD; REPEAT.  
 FT SIGNAL 1 25  
 FT CHAIN 26 1457  
 FT DOMAIN 26 752  
 FT TRANSMEM 753 774  
 FT DOMAIN 775 1457  
 FT DOMAIN 30 193  
 FT DOMAIN 208 276  
 FT DOMAIN 296 383  
 FT DOMAIN 392 473  
 FT DOMAIN 493 578  
 FT DOMAIN 596 679  
 FT DOMAIN 922 1158  
 FT DOMAIN 1218 1452  
 FT ACT\_SITE 1100 1100  
 FT ACT\_SITE 1394 1394  
 FT DISULFID 215 269  
 FT CARBOHYD 100 100  
 FT CARBOHYD 139 139  
 FT CARBOHYD 210 210  
 FT CARBOHYD 415 415  
 FT CARBOHYD 423 423  
 FT CARBOHYD 435 435  
 FT CARBOHYD 461 461  
 FT CARBOHYD 551 551

FT CARBOHYD 585 585 POTENTIAL.  
 FT CARBOHYD 589 589 POTENTIAL.  
 FT CARBOHYD 606 606 POTENTIAL.  
 FT CARBOHYD 689 689 POTENTIAL.  
 SQ SEQUENCE 1457 AA; 164185 MW; AEB67CCA CRC32;  
  
 Query Match 80.0%; Score 32; DB 1; Length 1457;  
 Best Local Similarity 77.8%; Pred. No. 28;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 VKALRVRR 9  
 DB 1132 VKALRSRR 1140  
  
 RESULT 6  
 TRPE\_SYNY3 STANDARD; PRT; 218 AA.  
 ID TRPE\_SYNY3  
 AC P74435;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE N-(5'-PHOSPHORIBOSYL)ANTHRANILATE ISOMERASE (EC 5.3.1.24) (PRAI).  
 GN TRPE OR SL0356.  
 OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).  
 OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 9761201.  
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,  
 RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,  
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K.,  
 RA OKUMURA S., SHIMO S., TAKEUCHI C., WADA T., MATANABE A.,  
 RA YAMADA M., YASUDA M., TABATA S.:  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions."  
 RL DNA RES. 3:109-136(1996).  
 CC -1- CATALYTIC ACTIVITY: N-(5'-PHOSPHORIBOSYL)-ANTHRANILATE =  
 CC 1-(2-CARBOXYPHENYLAMINO)-1-DEOXY-RIBULOSE 5-PHOSPHATE.  
 CC -1- PATHWAY: THIRD STEP IN TRYPTOPHAN BIOSYNTHESIS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D90915; G1653624; -  
 DR PFAM; PF00697; PRAI; 1.  
 DR HSP; Q56320; INSU.  
 KW ISOMERASE; TRYPTOPHAN BIOSYNTHESIS.  
 SQ SEQUENCE 218 AA; 23881 MW; 2AFB9F3F CRC32;  
  
 Query Match 80.0%; Score 32; DB 1; Length 218;  
 Best Local Similarity 75.0%; Pred. No. 3.5;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 VKALRVRR 8  
 DB 107 VKALRLRR 114  
  
 RESULT 7  
 TRLA\_ECOLI STANDARD; PRT; 241 AA.  
 ID TRLA\_ECOLI  
 AC Q00186;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)

SQ SEQUENCE 373 AA; 39666 MW; 264B834F CRC32;  
 Query Match 97.5%; Score 39; DB 1; Length 373;  
 Best Local Similarity 88.9%; Pred. No. 0.16;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 VKALVRRL 9  
 Db 85 IKALVRRL 93  
 RESULT 4  
 ID PTPK\_HUMAN STANDARD; PRT; 1439 AA.  
 AC Q15262; Q14763;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE PROTEIN-TYROSINE PHOSPHATASE KAPPA PRECURSOR (EC 3.1.3.48) (R-PTP-KAPPA).  
 GN PTPRK OR PTPK.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96279245.  
 RA FUCHS M., MUELLER T., LERCH M., ULLRICH A.;  
 RT "Association of human protein-tyrosine phosphatase kappa with members of the armadillo family."  
 RL J. BIOL. CHEM. 271:16712-16719(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-FORESKIN.  
 RA MEDLINE; 97199372.  
 RA YANG Y., GIL M.C., CHOI E.Y., PARK S.H., PYUN K.H., HA H.;  
 RT "Molecular cloning and chromosomal localization of a human gene homologous to the murine R-PTP-kappa, a receptor-type protein tyrosine phosphatase."  
 RL GENE 186:77-82(1997).  
 CC -1- FUNCTION: REGULATION OF PROCESSES INVOLVING CELL CONTACT AND ADHESION SUCH AS GROWTH CONTROL, TUMOR INVASION, AND METASTASIS. FORMS COMPLEXES WITH BETA-CATENIN AND GAMMA-CATENIN/PLAKOGLOBIN. BETA-CATENIN MAY BE A SUBSTRATE FOR THE CATALYTIC ACTIVITY OF PTP-KAPPA.  
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2O) = PROTEIN TYROSINE + ORTHOPHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN; AT ADHERENS JUNCTIONS.  
 CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN LUNG, BRAIN AND COLON; LESS IN LIVER, PANCREAS, STOMACH, KIDNEY, PLACENTA AND MAMMARY CARCINOMA.  
 CC -1- PTM: THIS PROTEIN UNDERGOES PROTEOLYTIC PROCESSING.  
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IG-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; Z70660; E234080; -;  
 DR EMBL; L77886; G1479976; -;  
 DR MIM; 602545; -;  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 2.  
 DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 2.  
 DR PROSITE; PS50055; TYR\_PHOSPHATASE\_PTP; 2.  
 DR PROSITE; PS00740; MAM\_1; 1.

DR PROSITE; PS50060; MAM\_2; 1.  
 DR PFAM; PF00041; fn3; 2.  
 DR PFAM; PF00047; ig; 1.  
 DR PFAM; PF00102; Y.Phosphatase; 2.  
 DR PFAM; PF00629; MAM; 1.  
 DR HSSP; P28827; IRPM.  
 KW HYDROLASE; RECEPTOR; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
 KM IMMUNOGLOBULIN FOLD; REPEAT.  
 FT SIGNAL 1 26  
 FT CHAIN 27 1439  
 FT DOMAIN 27 752  
 FT TRANSMEM 753 774  
 FT DOMAIN 775 1439  
 FT DOMAIN 31 194  
 FT DOMAIN 209 277  
 FT DOMAIN 297 384  
 FT DOMAIN 393 474  
 FT DOMAIN 494 579  
 FT DOMAIN 597 680  
 FT DOMAIN 910 1141  
 FT DOMAIN 1202 1433  
 FT ACT\_SITE 1082 1082  
 FT ACT\_SITE 1376 1376  
 FT ACT\_SITE 216 270  
 FT DISULFID 216 644  
 FT SITE 641 101  
 FT CARBOHYD 101 140  
 FT CARBOHYD 140 211  
 FT CARBOHYD 211 416  
 FT CARBOHYD 416 424  
 FT CARBOHYD 424 436  
 FT CARBOHYD 436 462  
 FT CARBOHYD 462 552  
 FT CARBOHYD 552 585  
 FT CARBOHYD 585 590  
 FT CARBOHYD 590 607  
 FT CARBOHYD 607 690  
 FT CARBOHYD 690 9  
 FT CONFLICT 9 158  
 FT CONFLICT 158 284  
 FT CONFLICT 284 422  
 FT CONFLICT 422 422  
 FT CONFLICT 422 674  
 FT CONFLICT 674 715  
 FT CONFLICT 715 732  
 FT CONFLICT 732 1366  
 FT CONFLICT 1366 1366  
 SQ SEQUENCE 1439 AA; 162087 MW; 442132E0 CRC32;  
 Query Match 80.0%; Score 32; DB 1; Length 1439;  
 Best Local Similarity 77.8%; Pred. No. 27;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 VKALVRRL 9  
 Db 1114 VKALRSRI 1122  
 RESULT 5  
 ID PTPK\_MOUSE STANDARD; PRT; 1457 AA.  
 AC P35822;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE PROTEIN-TYROSINE PHOSPHATASE KAPPA PRECURSOR (EC 3.1.3.48) (R-PTP-KAPPA).  
 GN PTPRK OR PTPK.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCURIONATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RI; TISSUE-BRAIN;

```

CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M86511; G180021; -.
DR EMBL; X13334; G29741; -.
DR EMBL; X06882; G312399; -.
DR PIR; A27637; TDHUM4.
DR PIR; A30555; A30555.
DR MIM; 158120; -.
DR PFAM; PF00560; LRR; 3.
DR KW ANTIGEN; GLYCOPROTEIN; GPI-ANCHOR; SIGNAL.
FT SIGNAL
FT CHAIN 1 19
FT CHAIN 20 375
FT CHAIN 20 367
FT CARBOHYD 37 37
FT CARBOHYD 151 151
FT CARBOHYD 282 282
FT CARBOHYD 323 323
FT CONFLICT 77 77
FT CONFLICT 187 187
FT CONFLICT 187 187
FT SEQUENCE 375 AA; 40136 MM; 98478838 CRC32;
SO
Oy 1 VKALVRRL 9
Db 86 VKALVRRL 94
Query Match 100.0%; Score 40; DB 1; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.096;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 2
ID CD14_RABIT STANDARD: PRT: 372 AA.
AC Q28680.
DC 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE MONOCYTE DIFFERENTIATION ANTIGEN CD14 PRECURSOR.
GN CD14.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUAROTTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE: 92268819.
CC LEE J.D., KATO K., TOBIAS P.S., KIRKLAND T.N., ULEVITCH R.J.;
CC "transfection of CD14 into 702/3 cells dramatically enhances the
CC RT sensitivity to complexes of lipopolysaccharide (LPS) and LPS binding
CC RT protein."
CC J. EXP. MED. 175:1697-1705(1992).
CC [2]
CC SEQUENCE FROM N.A.
CC TISSUE=MYELOID;
CC ISHIDA T., SEOGUCHI M., MATSURA K., YASUNORI H., AKIZUKI S.,
CC YAMAMOTO S.;
CC SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: SERVES AS AN LPS RECEPTOR CONTROLLING CELL ACTIVATION
CC UNDER PHYSIOLOGICAL CONDITIONS. WHEN LPS BINDS TO CD14 THE CELLS
CC BECOME ACTIVATED AND RELEASE CYTOKINES AND UP-REGULATE CELL
CC SURFACE MOLECULES, INCLUDING ADHESION MOLECULES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the swiss institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC
DR EMBL; M85233; G164865; -.
DR EMBL; D16545; D1022626; -.
DR PFAM; PF00560; LRR; 3.
KW ANTIGEN; GLYCOPROTEIN; GPI-ANCHOR; SIGNAL.
FT CHAIN 1 17 POTENTIAL.
FT CARBOHYD 18 372 CD14 DIFFERENTIATION ANTIGEN.
FT CARBOHYD 37 37 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 279 279 POTENTIAL.
FT CARBOHYD 320 320 POTENTIAL.
SQ SEQUENCE 372 AA; 39992 MW; CECDDFA8 CRC32;

Query Match 100.0%; Score 40; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 0.095;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKALRVRL 9
Db 86 VKALRVRL 94

RESULT 3
CD14_BOVIN STANDARD; PRT; 373 AA.
AC 095122;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE MONOCYTE DIFFERENTIATION ANTIGEN CD14 PRECURSOR (LPS RECEPTOR) (LPS-R)
DE (MELOID CELL-SPECIFIC LEUCINE-RICH GLYCOPROTEIN).
GN CD14.
OS BOS TAURUS (BOVINE).
CC EUGAROTIA; METAOTIA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
CC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HOLSTEIN;
RX MEDLINE; 97445486.
RA IKEDA A., TAKATA M., TANIGUCHI T., SEKIKAWA K.;
RT "Molecular cloning of bovine CD14 gene.";
RL J. VET. MED. SCI. 59:715-719(1997).
RN [2]
RP SEQUENCE OF 179-278 FROM N.A.
RX MEDLINE; 96209874.
RA DIAMOND G., RUSSELL J.P., BEVINS C.L.;
RT "Inducible expression of an antiobiotic peptide gene in
RT 11popolysaccharide-challenged tracheal epithelial cells.";
RC PROC. NATL. ACAD. SCI. U.S.A. 93:5156-5160(1996).
CC -I- FUNCTION: SERVES AS AN LPS RECEPTOR CONTROLLING CELL ACTIVATION
CC UNDER PHYSIOLOGICAL CONDITIONS. WHEN LPS BINDS TO CD14 THE CELLS
CC BECOME ACTIVATED AND RELEASE CYTOKINES AND UP-REGULATE CELL
CC SURFACE MOLECULES, INCLUDING ADHESION MOLECULES (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (BY
CC SIMILARITY).
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D84509; D1022366; -.
KW EMBL; U48356; G1532131; -.
KW ANTIGEN; GLYCOPROTEIN; GPI-ANCHOR; SIGNAL.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 1 373 MONOCYTE DIFFERENTIATION ANTIGEN CD14.
FT CARBOHYD 38 38 POTENTIAL.
FT CARBOHYD 150 150 POTENTIAL.
FT CARBOHYD 280 280 POTENTIAL.

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:28 ; Search time 71.09 Seconds  
(without alignments)  
3.579 Million cell updates/sec

Title: US-09-124-280a-21

Perfect score: 40

Sequence: 1 VKALRYRRL 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database: SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	375	1	CD14_HUMAN
2	40	100.0	372	1	CD14_RABIT
3	39	97.5	373	1	CD14_BOVIN
4	32	80.0	1439	1	PTPK_HUMAN
5	32	80.0	1457	1	PTPK_MOUSE
6	32	80.0	218	1	TRPF_SYNY3
7	31	77.5	241	1	TRL4_ECOLI
8	31	77.5	241	1	TRL5_ECOLI
9	30	75.0	451	1	SR54_METUA
10	29	72.5	1128	1	DNBI_EBY
11	29	72.5	513	1	PHSL_DESBA
12	29	72.5	466	1	RBL_DRORE
13	29	72.5	443	1	RBL_VERBO
14	29	72.5	486	1	UXUB_ECOLI
15	29	72.5	345	1	VG59_HSV1
16	29	72.5	448	1	YMI9_CABEL
17	28	70.0	1304	1	CD43_HUMAN
18	28	70.0	440	1	DNAI1_THEMA
19	28	70.0	205	1	FADD_MOUSE
20	28	70.0	305	1	FMT_THERH
21	28	70.0	236	1	GT6_SCHMA
22	28	70.0	1018	1	IDE_HUMAN
23	28	70.0	1019	1	IDE_RAT
24	28	70.0	843	1	MVPA_DICDI
25	28	70.0	365	1	PHEA_PSEST
26	28	70.0	756	1	PR11_HUMAN
27	28	70.0	614	1	RPB1_METYA
28	28	70.0	507	1	RRP1_RINDK
29	28	70.0	507	1	RRP1_RINDK
30	28	70.0	1395	1	SP41_YEAST
31	28	70.0	299	1	VY_RINDR
32	27	67.5	492	1	AMID_MORCA
33	27	67.5	513	1	C773_SOYBN
34	27	67.5	626	1	DPO2_SUISO
35	27	67.5	781	1	DPOL_ARCFU
36	27	67.5	2907	1	FBN2_MOUSE
37	27	67.5	783	1	FURI_XENIA
38	27	67.5	339	1	GPRH_HUMAN
39	27	67.5	669	1	ILVB_SCHPO
40	27	67.5	647	1	KMIL_CHICK
41	27	67.5	781	1	KRAF_CHICK
42	27	67.5	648	1	KRAF_HUMAN
43	27	67.5	648	1	KRAF_RAT

## ALIGNMENTS

RESULT	1	STANDARD	PRT	375 AA.
CD14_HUMAN				
ID	CD14_HUMAN			
AC	P08571			
DT	01-AUG-1988 (REL. 08, CREATED)			
DT	01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)			
DE	MONOCYTE DIFFERENTIATION ANTIGEN CD14 PRECURSOR (LPS RECEPTOR) (LPS-R)			
DE	(MYELOID CELL-SPECIFIC LEUCINE-RICH GLYCOPROTEIN).			
GN	CD14.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-LYMPHOCYTES;			
RX	MEDLINE; 88234022.			
RA	FERRERO E., GOYERT S.M.;			
RT	"Nucleotide sequence of the gene encoding the monocyte			
RT	differentiation antigen, CD14."			
RL	NUCLEIC ACIDS RES. 16:4173-4173(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 88258060.			
RA	HAZIOR A., CHEN S., FERRERO E., LOW M.G., SILBER R., GOYERT S.M.;			
RT	"The monocyte differentiation antigen, CD14, is anchored to the cell			
RT	membrane by a phosphatidylinositol linkage."			
RL	J. IMMUNOL. 141:547-552(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-MACROPHAGE;			
RX	MEDLINE; 89287330.			
RA	SETOUCHI M., NASU N., YOSHIDA S., HIGUCHI Y., AKIZUKI S.,			
RA	YAMAMOTO S.;			
RT	"Mouse and human CD14 (myeloid cell-specific leucine-rich			
RT	glycoprotein) primary structure deduced from cDNA clones."			
RL	BIOCHIM. BIOPHYS. ACTA 1008:213-222(1989).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 89088340.			
RA	SIMMONS D.L., TAN S., TENEN D.G., NICHOLSON-WELLER A., SEED B.;			
RT	"Monocyte antigen CD14 is a phospholipid anchored membrane protein."			
RL	BLOOD 73:284-289(1989).			
RN	[5]			
RP	SEQUENCE OF 362-367.			
RX	MEDLINE; 89384684.			
RA	BAJIL V., BAUDYS M., HILGERT I., STEFANOVA I., LOW M.G., ZBROZEK J.,			
RA	HOREJSEI V.;			
RT	"Structural relationship between the soluble and membrane-bound forms			
RT	of human monocyte surface glycoprotein CD14."			
RL	MOL. IMMUNOL. 26:657-662(1989).			
CC	-1- FUNCTION: SERVES AS AN LPS RECEPTOR CONTROLLING CELL ACTIVATION			
CC	UNDER PHYSIOLOGICAL CONDITIONS. WHEN LPS BINDS TO CD14 THE CELLS			
CC	BECOME ACTIVATED AND RELEASE CYTOKINES AND UP-REGULATE CELL			
CC	SURFACE MOLECULES, INCLUDING ADHESION MOLECULES.			
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED STRONGLY ON THE SURFACE OF MONOCYTES			
CC	AND WEAKLY ON THE SURFACE OF GRANULOCYTES; ALSO EXPRESSED BY MOST			
CC	TISSUE MACROPHAGES.			
CC	-1- DATABASE: NAME=PROV; NOTE=CD guide CD14 entry;			
CC	WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd14.htm.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			

**THIS PAGE BLANK (USPTO)**

C:Keywords: carbon-oxygen lyase; hydro-lyase; NAD; oxidoreductase

Query Match 72.5%; Score 29; DB 2; Length 486;  
Best Local Similarity 75.0%; Pred. No. 84;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRR 8  
|:|||||  
DB 166 VEALRLRR 173

## RESULT 14

S18587  
hypothetical protein 2 - Thibacillus ferrooxidans  
C:Species: Thibacillus ferrooxidans  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Sep-1997  
C:Accession: S18587  
R:Inoue, C.; Sugawara, K.; Kusano, T.  
Mol. Microbiol. 5, 2707-2718, 1991  
A:Title: The merr regulatory gene in Thibacillus ferrooxidans is spaced apart from the  
A:Reference number: S18584; MUID:92140035  
A:Accession: S18587  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-138 <IMO>  
A:Cross-references: EMBL:X57326; NID:948150; PID:948154

Query Match 72.5%; Score 29; DB 2; Length 138;  
Best Local Similarity 62.5%; Pred. No. 23;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRR 8  
|:|||||  
DB 125 VQAMRIRR 132

## RESULT 15

S31129  
hypothetical protein F59B2.9 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Sep-1997  
C:Accession: S31129  
R:Sturston, J.; Du, Z.; Thomas, K.; Wilson, R.; Hillier, L.; Staden, R.; Halloran, N.; G.  
awkins, T.; Ainscough, R.; Waterston, R.  
submitted to the EMBL Data Library, November 1991  
A:Description: The C. elegans sequencing project: A beginning.  
A:Reference number: S31122  
A:Accession: S31129  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-448 <SUL>  
A:Cross-references: EMBL:Z11505; NID:96718; PID:96726  
C:Genetics:  
A:Introns: 25/2; 98/2; 227/3

Query Match 72.5%; Score 29; DB 2; Length 448;  
Best Local Similarity 66.7%; Pred. No. 77;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VKALRVRR 9  
|:|||||  
DB 226 IKALRVREL 234

Search completed: September 7, 1999, 23:22:32  
Job time: 939 sec



Matches 5: Conservative 3: Mismatches 1: Indels 0: Gaps 0:  
OY 1 VKALVRRL 9  
: ||| :  
Db 167 LKALEIRRI 175

RESULT 10  
Q0BE47  
DNA-binding protein - human herpesvirus 4 (strain B95-8)  
C:Species: human herpesvirus 4, Epstein-Barr virus  
C>Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 05-Sep-1997  
C:Accession: A43045; A03791; S33057  
R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrrell, B.G.  
Mol. Biol. Med. 1, 21-45, 1983  
A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus  
A:Reference number: A93065; MUID:85035713  
A:Accession: A43045  
A:Molecule type: DNA  
A:Residues: 1-1128 <BAN>  
A:Cross-references: EMBL:V0155; NID:959074; PID:q134916  
R:Beer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H  
Nature 310, 207-211, 1984  
A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.  
A:Reference number: A03794; MUID:84270667  
A:Contents: annotation; protein coding region  
C:Superfamily: herpesvirus DNA-binding protein  
C:Keywords: DNA binding

Query Match 72.5% Score 29; DB 1; Length 1128;  
Best Local Similarity 75.0% Pred. No. 2e+02;  
Matches 6: Conservative 2: Mismatches 0: Indels 0: Gaps 0:  
OY 2 KALVRRL 9  
: ||||| :  
Db 223 QALVRRV 230

RESULT 11  
MMBET7  
38.1K membrane protein - Ictalurid herpesvirus 1 (strain auburn 1)  
C:Species: Ictalurid herpesvirus 1  
A:Note: host Ictalurus punctatus (channel catfish)  
C>Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 05-Sep-1997  
C:Accession: F36792  
R:Davidson, A.J.  
Submitted to Genbank, January 1992  
A:Description: Channel catfish virus: a new type of herpesvirus.  
A:Reference number: A36804  
A:Accession: F36792  
A:Molecule type: DNA  
A:Residues: 1-345 <DAV>  
A:Cross-references: GB:M75136; NID:g331209; PID:g331269  
R:Davidson, A.J.  
Virology 186, 9-14, 1992  
A:Title: Channel catfish virus: a new type of herpesvirus.  
A:Reference number: A39447; MUID:92087490  
A:Contents: annotation  
A:Note: neither amino acid nor nucleotide sequence is given  
C:Genetics:  
A:Gene: 59  
C:Superfamily: Ictalurid herpesvirus 38.1K membrane protein  
C:Keywords: transmembrane protein  
F:46-63/Domain: transmembrane #status predicted <TM1>  
F:101-118/Domain: transmembrane #status predicted <TM2>  
F:147-165/Domain: transmembrane #status predicted <TM3>  
F:265-286/Domain: transmembrane #status predicted <TM4>

Query Match 72.5% Score 29; DB 1; Length 345;  
Best Local Similarity 66.7% Pred. No. 59;  
Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

OY 1 VKALVRRL 9  
: ||||| :  
Db 286 IRALRRRL 294

RESULT 12  
A30610  
chloroplast rbc L gene protein - grape  
C:Species: vitis sp. (grape)  
C>Date: 30-Sep-1993 #sequence\_revision 19-Oct-1995 #text\_change 22-May-1998  
C:Accession: A30610  
R:Huang, Y.; Ma, C.; Li, C.L.; Wu, N.H.  
Acta Bot. Sin. 36, 444-451, 1994  
A:Title: Structure of the chloroplast rbc L gene from grape (vitis vinifera).  
A:Reference number: A30610  
A:Accession: A30610  
A:Molecule type: DNA  
A:Residues: 1-475 <HUA>  
C:Genetics:  
A:Gene: rbc L  
C:Superfamily: ribulose-bisphosphate carboxylase large chain  
C:Keywords: chloroplast; magnesium  
F:175,334/Active site: Lys (ribulose-bisphosphate-binding) #status predicted  
F:201/Binding site: carbon dioxide (Lys) (covalent) (by Rubisco activase) #status pre  
F:203/Binding site: magnesium (Asp) #status predicted

Query Match 72.5% Score 29; DB 2; Length 475;  
Best Local Similarity 77.8% Pred. No. 82;  
Matches 7: Conservative 0: Mismatches 2: Indels 0: Gaps 0:  
OY 1 VKALVRRL 9  
: ||||| :  
Db 127 VKALRLRL 135

RESULT 13  
S56548  
fructuronate reductase (EC 1.1.1.57) uxub - Escherichia coli  
N:Alternate names: D-mannonate oxidoreductase  
C:Species: Escherichia coli  
C>Date: 28-Oct-1995 #sequence\_revision 13-Mar-1997 #text\_change 01-Feb-1999  
C:Accession: S56548; E65246  
R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.  
Nucleic Acids Res. 23, 2105-2119, 1995  
A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from  
A:Reference number: S56514; MUID:95394362  
A:Accession: S56548  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-486 <BUR>  
A:Cross-references: EMBL:U14003; NID:g1263172; PID:g537164  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426517  
A:Accession: E65246  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-486 <BLAT>  
A:Cross-references: GB:AE00503; GB:U00096; NID:q1790777; PID:q1790779; UMGF:b4323  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: uxub  
C:Function:  
A:Description: reversible reaction of D-mannonate and NAD(+) to D-fructuronate and NA  
A:Pathway: D-gluconate catabolism  
A:Note: part of the interconversion of D-fructuronate to D-2-keto-3-deoxy-D-gluconate  
C:Superfamily: conserved hypothetical protein YEL070W



F:37,151,282,323/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 40; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. NO. 0.26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
| | | | | | | |  
Db 86 VKALRVRL 94

RESULT 2  
A:1326  
enantiomer-selective amidase - Rhodococcus sp.  
C:Species: Rhodococcus sp.  
C>Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 12-Sep-1997  
C:Accession: A41326  
R:Mayaux, J.F.; Cerbelaud, E.; Soubrier, F.; Yeh, P.; Blanche, F.; Petre, D.  
J. Bacteriol. 173, 6594-6704, 1991  
A:Title: Purification, cloning, and primary structure of a new enantiomer-selective amidase  
A:Reference number: A41326; MUID:92041549  
A:Accession: A41326  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-462 <MAY>  
A:Cross-references: GB:M74531; NID:g152051; PID:g152052  
C:Superfamily: indoleacetamide hydrolase

Query Match 82.5%; Score 33; DB 2; Length 462;  
Best Local Similarity 77.8%; Pred. NO. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
| | | | | | | |  
Db 347 IKALRVRL 355

RESULT 3  
A:8066  
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type kappa precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 21-Jan-1994 #sequence\_revision 01-Mar-1996 #text\_change 12-Feb-1999  
C:Accession: A48066  
R:Jiang, Y.P.; Wang, H.; D'Eustachio, P.; Musacchio, J.M.; Schlessinger, J.; Sap, J.  
Mol. Cell. Biol. 13, 2942-2951, 1993  
A:Title: Cloning and characterization of R-PTP-kappa, a new member of the receptor protein.  
A:Reference number: A48066; MUID:93233655  
A:Accession: A48066  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1457 <JIA>  
A:Cross-references: GB:L10106; NID:g293771; PID:g293772  
A:Experimental source: brain  
A>Note: sequence extracted from NCBI backbone (NCBIP:129880)  
C:Complete: after cleavage by a furin-like endopeptidase, the two chains remain associated  
C:Superfamily: protein-tyrosine-phosphatase, receptor type mu; fibronectin type III repeat  
-tyrosine-phosphatase homology  
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; receptor; trans  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-147/Product: protein-tyrosine-phosphatase, receptor type kappa, uncleaved #status  
F:29-641/Product: protein-tyrosine-phosphatase, receptor type kappa, chain 1 #status pre  
F:29-641,644-755/Domain: extracellular #status predicted <EXT>  
F:30-193/Domain: MAM homology <MAM>  
F:208-271/Domain: immunoglobulin homology <IGH>  
F:293-371/Domain: fibronectin type III repeat homology <FN3A>  
F:388-477/Domain: fibronectin type III repeat homology <FN3B>  
F:489-551/Domain: fibronectin type III repeat homology <FN3C>  
F:551-652/Domain: fibronectin type III repeat homology <FN3D>  
F:644-1457/Product: protein-tyrosine-phosphatase, receptor type kappa, chain 2 #status  
F:756-772/Domain: transmembrane #status predicted <TMN>

F:773-1457/Domain: intracellular #status predicted <INT>  
F:841-1457/Domain: leukocyte common antigen cytosolic domain homology <LAC>  
F:922-1148/Domain: protein-tyrosine-phosphatase homology <PTP1>  
F:1218-1442/Domain: protein-tyrosine-phosphatase homology <PTP2>  
F:100,139,210,415,423,435,461,551,585,589,606,689/Binding site: carbohydrate (Asn) (c  
F:215-269/Disulfide bonds: #status predicted  
F:643-644/Cleavage site: Arg-Glu (furin-like endopeptidase) #status experimental  
F:1100/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:1106/Binding site: substrate phosphate (Arg) #status predicted  
F:1394/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:1400/Binding site: substrate phosphate (Arg) #status predicted

Query Match 80.0%; Score 32; DB 1; Length 1457;  
Best Local Similarity 77.8%; Pred. NO. 56;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
| | | | | | | |  
Db 1132 VKALRSRI 1140

RESULT 4  
S76407  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
C:Accession: S76407  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
S.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76407  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-218 <KAN>  
A:Cross-references: EMBL:D90915; GB:AB001339; NID:g1653604; PID:d1019269; PID:g165362  
C:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: phosphoribosylanthranilate isomerase; trpF homology  
F:6-210/Domain: trpF homology <TRF>

Query Match 80.0%; Score 32; DB 2; Length 218;  
Best Local Similarity 75.0%; Pred. NO. 8.3;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 8  
| | | | | | | |  
Db 107 IKALRVRL 114

RESULT 5  
JC6312  
protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type - human  
C:Species: Homo sapiens (hmn)  
C>Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 12-Feb-1999  
C:Accession: JC6312  
R:Yang, Y.; Gill, M.C.; Choi, E.Y.; Park, S.H.; Pyun, K.H.; Ha, H.  
Gene 166, 77-82, 1997  
A:Title: Molecular cloning and chromosomal localization of a human gene homologous to  
A:Reference number: JC6312; MUID:97199372  
A:Accession: JC6312  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1440 <YAN>  
A:Cross-references: GB:L77886  
C:Superfamily: protein-tyrosine-phosphatase, receptor type mu; fibronectin type III r  
-tyrosine-phosphatase homology  
C:Keywords: phosphoric monoester hydrolase

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:53 ; Search time 105.14 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-21  
Perfect score: 40  
Sequence: 1 VKALRVRL 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	375	1	TDHUM4	monocyte surface g
2	33	82.5	462	2	A41326	enantiomer-selecti
3	32	80.0	1457	1	A48066	protein-tyrosine-p
4	32	80.0	218	2	S76407	hypothetical prote
5	32	80.0	1440	2	JC6312	protein-tyrosine-p
6	31	77.5	241	2	S23004	trial protein - Esc
7	30	75.0	451	2	E64312	signal recognition
8	30	75.0	2958	2	S64921	probable membrane
9	29	72.5	514	1	HQDVLB	cytochrome-c3 hydr
10	29	72.5	1138	1	OOBE47	DNA-binding protei
11	29	72.5	345	1	MMBER7	38 kR membrane pro
12	29	72.5	475	2	A30610	chloroplast rbc L
13	29	72.5	466	2	S65348	fructuronate reduc
14	29	72.5	488	2	S18587	hypothetical prote
15	29	72.5	448	2	S31129	hypothetical prote
16	29	72.5	286	2	S50855	neurotrophin-6 - s
17	28	70.0	1019	1	SNHUTN	insulinase (EC 3.4
18	28	70.0	1019	1	S29509	insulinase (EC 3.4
19	28	70.0	507	1	A43387	polymerase-associ
20	28	70.0	398	2	E70113	acetyl-CoA C-acety
21	28	70.0	512	2	I49552	protein-tyrosine k
22	28	70.0	478	2	F70322	glutamyl-tRNA (Gln
23	28	70.0	365	2	A44764	chorismate mutase
24	28	70.0	1304	2	A46546	leukocyte common a
25	28	70.0	237	2	S55614	hypothetical prote
26	28	70.0	507	2	JQ1929	phosphoprotein - r
27	28	70.0	299	2	JQ1930	V protein - rinder
28	28	70.0	305	2	B55228	methionyl-tRNA for
29	28	70.0	614	2	A47161	DNA-directed RNA p
30	28	70.0	310	2	H69465	hypothetical prote
31	28	70.0	1004	2	B69483	hypothetical prote
32	28	70.0	843	2	A47132	major vault protei
33	28	70.0	1435	2	S69632	regulatory protein
34	28	70.0	239	2	E70315	conserved hypotet
35	28	70.0	378	2	E71274	probable ATP-bindi
36	27	67.5	379	1	OKHUK1	protein kinase (EC
37	27	67.5	648	1	TVHUF6	protein kinase raf
38	27	67.5	648	1	TVTRF6	protein kinase raf
39	27	67.5	647	1	S00644	protein kinase raf

40	27	67.5	781	1	TVFDF	protein kinase Dra
41	27	67.5	2347	1	TVHUR5	kinase-related pro
42	27	67.5	514	1	QJMPB	photosystem II chl
43	27	67.5	626	2	S34120	DNA-directed DNA p
44	27	67.5	596	2	B41627	furin homolog 18 -
45	27	67.5	494	2	JC5919	potassium channel

## ALIGNMENTS

RESULT 1  
TDHUM4  
monocyte surface glycoprotein CD14 precursor - human  
N:Alternate names: monocyte differentiation antigen CD14; myeloid cell-specific leuci  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1990 #sequence\_revision 30-Jun-1990 #text\_change 05-Sep-1997  
C:Accession: A27637; S00917; S04968; P10108; A30555  
R:Haziot, A.; Chen, S.; Ferrero, E.; Low, M.G.; Silber, R.; Goyert, S.M.  
J. Immunol. 141, 547-552, 1988  
A:Title: The monocyte differentiation antigen, CD14, is anchored to the cell membrane  
A:Reference number: A27637; MOID:8828060  
A:Accession: A27637  
A:Molecule type: mRNA  
A:Residues: 1-375 <HAZ>  
Nucleic Acids Res. 16, 4173, 1988  
A:Title: Nucleotide sequence of the gene encoding the monocyte differentiation antige  
A:Reference number: S00917; MOID:88234022  
A:Accession: S00917  
A:Molecule type: DNA  
A:Residues: 1-375 <FER>  
A:Cross-references: EMBL:X06882; NID:g29736; PID:g312399  
R:Seoguchi, M.; Nasu, N.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
Biochim. Biophys. Acta 1008, 213-222, 1989  
A:Title: Mouse and human CD14 (myeloid cell-specific leucine-rich glycoprotein) prima  
A:Reference number: S04967; MOID:89287330  
A:Accession: S04968  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-186, 'C', 188-375 <SEN>  
A:Cross-references: EMBL:X13334; NID:g29740; PID:g29741  
R:Basil, V.; Baudys, M.; Hiltger, I.; Stefanova, I.; Low, M.G.; Zbrozek, J.; Horejst, M.  
Mol. Immunol. 26, 657-662, 1989  
A:Title: Structural relationship between the soluble and membrane-bound forms of huma  
A:Reference number: P10108; MOID:89384684  
A:Accession: P10108  
A:Molecule type: protein  
A:Residues: 362-367 <BAZ>  
R:Simmons, D.L.; Tan, S.; Tenen, D.G.; Nicholson-Weller, A.; Seed, B.  
Blood 73, 284-289, 1989  
A:Title: Monocyte antigen CD14 is a phospholipid anchored membrane protein.  
A:Reference number: A30555; MOID:89088540  
A:Accession: A30555  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-76, 'G', 78-186, 'C', 188-375 <SIM>  
C:Comment: This glycoprotein is expressed primarily on peripheral blood monocytes and  
ells expressing the protein on their surfaces.  
C:Comment: The protein is anchored to the cell membrane by a phosphatidylinositol lin  
e membrane-bound form.  
C:Genetics:  
A:Gene: GDB:CD14  
A:Cross-references: GDB:119759; OMIM:158120  
A:Map position: 5q31.1-5q31.1  
A:Insertions: 1/3  
C:Superfamily: monocyte surface glycoprotein CD14  
C:Keywords: glycoprotein; membrane protein; phosphatidylinositol linkage  
F:1-19/Domain: signal sequence #status predicted <Sig>  
F:20-367/Product: monocyte surface glycoprotein CD14 #status predicted <NAT>  
F:20-367/Product: monocyte surface glycoprotein CD14, soluble form #status predicted  
F:91-329/Region: 9-residue repeats ([LV]-X-X-L-X-[LVIV]-[SRAT]-X-[NAT])

ATTORNEY/AGENT INFORMATION:  
NAME: Curry, Daniel R.  
REGISTRATION NUMBER: 32,727  
REFERENCE/DOCKET NUMBER: A-324A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805/447-1000  
TELEFAX: 805/447-1090  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-10

Query Match 100.0%; Score 40; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
Db 67 VKALRVRL 75

RESULT 14  
US-08-484-397A-11  
Sequence 11, Application US/08484397A  
Patent No. 5869055  
GENERAL INFORMATION:  
APPLICANT: Juan, Shao-Chieh  
APPLICANT: Lichenstein, Henri S.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: AMGEN INC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: CA  
COUNTRY: US  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,397A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Curry, Daniel R.  
REGISTRATION NUMBER: 32,727  
REFERENCE/DOCKET NUMBER: A-324A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805/447-1000  
TELEFAX: 805/447-1090  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-11

Query Match 100.0%; Score 40; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
Db 67 VKALRVRL 75

RESULT 15  
US-08-484-397A-12  
Sequence 12, Application US/08484397A  
Patent No. 5869055  
GENERAL INFORMATION:  
APPLICANT: Juan, Shao-Chieh  
APPLICANT: Lichenstein, Henri S.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: AMGEN INC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: CA  
COUNTRY: US  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,397A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Curry, Daniel R.  
REGISTRATION NUMBER: 32,727  
REFERENCE/DOCKET NUMBER: A-324A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805/447-1000  
TELEFAX: 805/447-1090  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-12

Query Match 100.0%; Score 40; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
Db 67 VKALRVRL 75

Search completed: September 7, 1999, 23:07:43  
Job time: 1741 sec

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-7

Query Match 100.0%; Score 40; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
Db 67 VKALRVRL 75

RESULT 11  
US-08-484-397A-8

; Sequence 8, Application US/08484397A  
; Patent No. 5869055  
; GENERAL INFORMATION:  
; APPLICANT: Juan, Shao-Chieh  
; APPLICANT: Lichenstein, Henri S.  
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,397A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Curry, Daniel R.  
; REGISTRATION NUMBER: 32,727  
; REFERENCE/DOCKET NUMBER: A-324A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 805/447-1090  
; TELEFAX: 805/447-1090  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 334 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-484-397A-8

Query Match 100.0%; Score 40; DB 2; Length 334;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
Db 53 VKALRVRL 61

RESULT 12  
US-08-484-397A-9  
; Sequence 9, Application US/08484397A  
; Patent No. 5869055  
; GENERAL INFORMATION:  
; APPLICANT: Juan, Shao-Chieh  
; APPLICANT: Lichenstein, Henri S.  
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,397A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424

APPLICANT: Juan, Shao-Chieh  
APPLICANT: Lichenstein, Henri S.  
APPLICANT: Wright, Samuel D.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: CA  
COUNTRY: US  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,397A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Curry, Daniel R.  
REGISTRATION NUMBER: 32,727  
REFERENCE/DOCKET NUMBER: A-324A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805/447-1090  
TELEFAX: 805/447-1090  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-9

Query Match 100.0%; Score 40; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
Db 67 VKALRVRL 75

RESULT 13  
US-08-484-397A-10  
; Sequence 10, Application US/08484397A  
; Patent No. 5869055  
; GENERAL INFORMATION:  
; APPLICANT: Juan, Shao-Chieh  
; APPLICANT: Lichenstein, Henri S.  
; APPLICANT: Wright, Samuel D.  
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,397A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424

OY 1 VKALRVRL 9  
Db 67 VKALRVRL 75

RESULT 8  
US-08-484-397A-5  
; Sequence 5, Application US/08484397A  
; Patent No. 5869055  
; GENERAL INFORMATION:  
; APPLICANT: Juan, Shao-Chieh  
; APPLICANT: Lichenstein, Henri S.  
; APPLICANT: Wright, Samuel D.  
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,397A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Curry, Daniel R.  
; REGISTRATION NUMBER: 32,727  
; REFERENCE/DOCKET NUMBER: A-324A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 805/447-1000  
; TELEFAX: 805/447-1090  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 348 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-484-397A-5

Query Match 100.0%; Score 40; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
Db 67 VKALRVRL 75

RESULT 9  
US-08-484-397A-6  
; Sequence 6, Application US/08484397A  
; Patent No. 5869055  
; GENERAL INFORMATION:  
; APPLICANT: Juan, Shao-Chieh  
; APPLICANT: Lichenstein, Henri S.  
; APPLICANT: Wright, Samuel D.  
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: CA

COUNTRY: US  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,397A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Curry, Daniel R.  
REGISTRATION NUMBER: 32,727  
REFERENCE/DOCKET NUMBER: A-324A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805/447-1090  
TELEFAX: 805/447-1090  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-6

Query Match 100.0%; Score 40; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
Db 67 VKALRVRL 75

RESULT 10  
US-08-484-397A-7  
; Sequence 7, Application US/08484397A  
; Patent No. 5869055  
; GENERAL INFORMATION:  
; APPLICANT: Juan, Shao-Chieh  
; APPLICANT: Lichenstein, Henri S.  
; APPLICANT: Wright, Samuel D.  
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,397A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Curry, Daniel R.  
; REGISTRATION NUMBER: 32,727  
; REFERENCE/DOCKET NUMBER: A-324A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 805/447-1090  
; TELEFAX: 805/447-1090  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 348 amino acids

GENERAL INFORMATION:  
APPLICANT: Juan, Shao-Chieh  
APPLICANT: Lichenstein, Henri S.  
APPLICANT: Wright, Samuel D.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: CA  
COUNTRY: US  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,397A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Curry, Daniel R.  
REGISTRATION NUMBER: 32,727  
REFERENCE/DOCKET NUMBER: A-324A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805/447-1090  
TELEFAX: 805/447-1090  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-2

Query Match 100.0%; Score 40; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
DB 67 VKALRVRL 75

RESULT 6  
US-08-484-397A-3  
Sequence 3, Application US/08484397A  
Patent No. 5869055  
GENERAL INFORMATION:  
APPLICANT: Juan, Shao-Chieh  
APPLICANT: Lichenstein, Henri S.  
APPLICANT: Wright, Samuel D.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: CA  
COUNTRY: US  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,397A  
FILING DATE: 07-JUN-1995

CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Curry, Daniel R.  
REGISTRATION NUMBER: 32,727  
REFERENCE/DOCKET NUMBER: A-324A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805/447-1090  
TELEFAX: 805/447-1090  
INFORMATION FOR SEQ. ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-3

Query Match 100.0%; Score 40; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
DB 67 VKALRVRL 75

RESULT 7  
US-08-484-397A-4  
Sequence 4, Application US/08484397A  
Patent No. 5869055  
GENERAL INFORMATION:  
APPLICANT: Juan, Shao-Chieh  
APPLICANT: Lichenstein, Henri S.  
APPLICANT: Wright, Samuel D.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: CA  
COUNTRY: US  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,397A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Curry, Daniel R.  
REGISTRATION NUMBER: 32,727  
REFERENCE/DOCKET NUMBER: A-324A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805/447-1090  
TELEFAX: 805/447-1090  
INFORMATION FOR SEQ. ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-397A-4

Query Match 100.0%; Score 40; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSEE: Patent Counsel  
STREET: 10666 No. 5705398th Torrey Pines Road, TPC 8  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/205,719  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: SCR1201P: TSP1324.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-205-719-2

Query Match 100.0%; Score 40; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALVRRL 9  
Db 86 VKALVRRL 94

RESULT 3  
US-08-366-953A-45  
Sequence 45, Application US/08366953A  
Patent No. 5765593  
GENERAL INFORMATION:  
APPLICANT: Lichenstein, Henri S.  
APPLICANT: Wright, Samuel D.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 PEPTIDES  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: 1840 Dehavenland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1720  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/366,953A  
FILING DATE: 30-DEC-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Cook Ph.D., Robert R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-324  
INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-366-953A-45

Query Match 100.0%; Score 40; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALVRRL 9  
Db 67 VKALVRRL 75

RESULT 4  
US-08-456-112B-21  
Sequence 21, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-21

Query Match 100.0%; Score 40; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.8e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALVRRL 9  
Db 1 VKALVRRL 9

RESULT 5  
US-08-484-397A-2  
Sequence 2, Application US/08484397A  
Patent No. 5869055

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:42 ; Search time 82.37 Seconds  
(without alignments)  
1.078 Million cell updates/sec

Title: US-09-124-280A-21  
Perfect score: 40  
Sequence: 1 VKALRVRL 9

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database:

Issued Patents.AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/PCFUS9.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	1	US-08-097-830E-21
2	40	100.0	375	1	US-08-205-719-2
3	40	100.0	348	2	US-08-366-953A-45
4	40	100.0	9	2	US-08-456-112B-21
5	40	100.0	348	2	US-08-484-397A-2
6	40	100.0	348	2	US-08-484-397A-3
7	40	100.0	348	2	US-08-484-397A-4
8	40	100.0	348	2	US-08-484-397A-5
9	40	100.0	348	2	US-08-484-397A-6
10	40	100.0	348	2	US-08-484-397A-7
11	40	100.0	334	2	US-08-484-397A-8
12	40	100.0	152	2	US-08-484-397A-9
13	40	100.0	152	2	US-08-484-397A-10
14	40	100.0	152	2	US-08-484-397A-11
15	40	100.0	152	2	US-08-484-397A-12
16	40	100.0	152	2	US-08-484-397A-13
17	40	100.0	152	2	US-08-484-397A-14
18	40	100.0	138	2	US-08-484-397A-15
19	40	100.0	348	2	US-08-484-397A-16
20	40	100.0	348	2	US-08-484-397A-17
21	33	82.5	461	1	US-07-796-361A-13
22	33	82.5	461	1	US-07-612-673-4
23	33	82.5	462	2	US-08-539-666-4
24	33	82.5	462	2	US-08-652-971-3
25	32	80.0	1457	2	US-08-449-644-1
26	32	80.0	1439	2	US-08-449-644-2
27	32	80.0	1439	2	US-08-449-644-3
28	32	80.0	1439	2	US-08-449-644-4
29	27	67.5	257	1	US-08-077-256-1
30	27	67.5	648	1	US-08-276-151-2
31	27	67.5	197	1	US-08-276-151-3
32	27	67.5	648	1	US-08-185-282-1
33	27	67.5	648	1	US-08-185-282-2
34	27	67.5	648	1	US-08-185-282-3
35	27	67.5	648	1	US-08-185-282-4
36	27	67.5	648	1	US-08-185-282-5
37	27	67.5	648	1	US-08-185-282-12
38	27	67.5	1237	1	US-08-241-853-2
39	27	67.5	494	1	US-08-464-340A-4

40	27	67.5	257	1	US-08-259-672-1	Sequence 1, Appl
41	27	67.5	149	1	US-08-259-672-6	Sequence 6, Appl
42	27	67.5	99	1	US-08-259-672-7	Sequence 7, Appl
43	27	67.5	81	1	US-08-259-672-8	Sequence 8, Appl
44	27	67.5	339	2	US-08-153-848-44	Sequence 44, Appl
45	27	67.5	574	2	US-08-309-512-9	Sequence 9, Appl

## ALIGNMENTS

```
RESULT 1
US-08-097-830E-21
; Sequence 21, Application US/08097830E
; Patent No. 5652211
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Peptides For Neutralizing The
; TITLE OF INVENTION: Toxicity of Lipid A
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-097-830E-21

Query Match 100.0%; Score 40; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9
Db 1 VKALRVRL 9

RESULT 2
US-08-205-719-2
; Sequence 2, Application US/08205719
; Patent No. 5705398
; GENERAL INFORMATION:
; APPLICANT: Mintz, D. N.
; APPLICANT: Tobias, P. S.
; APPLICANT: Ulevitch, R. J.
; TITLE OF INVENTION: SYSTEM AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: THERAPEUTIC INHIBITORS OF LPS-MEDIATED SEPSIS
```

Query Match 100.0%; Score 40; DB 1; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.32;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 |||||  
 DB 67 VKALRVRL 75

## RESULT 14

W00477  
 ID W00477 standard; protein; 152 AA.  
 AC W00477;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative (truncated).  
 KW CD14; treatment; prevention; inflammation; sepsis; liver failure;  
 KW adult respiratory distress syndrome; acute pancreatitis;  
 KW inflammatory bowel disease; graft versus host disease; bone marrow;  
 KW transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 PN WC9620957-A1.  
 PD 11-JUL-1996.  
 PE 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PA (AMGE-) AMGEN INC.  
 PA (UYRO-) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 PI WPI: 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 49-50; 79pp; English.  
 CC Derivatives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 152 AA;

Query Match 100.0%; Score 40; DB 1; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.32;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 |||||  
 DB 67 VKALRVRL 75

## RESULT 15

W00478  
 ID W00478 standard; protein; 152 AA.  
 AC W00478;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative (truncated).  
 KW CD14; treatment; prevention; inflammation; sepsis; liver failure;  
 KW adult respiratory distress syndrome; acute pancreatitis;  
 KW inflammatory bowel disease; graft versus host disease; bone marrow;  
 KW transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 PN WC9620957-A1.  
 PD 11-JUL-1996.  
 PE 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.

PA (AMGE-) AMGEN INC.  
 PA (UYRO-) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 PI WPI: 96-333944/33.

PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 50; 79pp; English.  
 CC Derivatives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 152 AA;

Query Match 100.0%; Score 40; DB 1; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.32;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 |||||  
 DB 67 VKALRVRL 75

Search completed: September 7, 1999, 22:49:51  
 Job time: 7872 sec

CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure;  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. . They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 348 AA;

Query Match 100.0%; Score 40; DB 1; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75

RESULT 11  
 W00474  
 ID W00474 standard; protein; 318 AA.  
 AC W00474;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative (N-terminal deletion).  
 KW adult respiratory distress syndrome; sepsis; liver failure;  
 KW CD14; treatment; prevention; inflammation; sepsis; acute pancreatitis;  
 KW inflammatory bowel disease; graft versus host disease; bone marrow;  
 KW transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1..18  
 FT disulfide\_bond 154..184  
 FT disulfide\_bond 208..239  
 FT disulfide\_bond 273..319  
 PN W09620957-A1.  
 PD 11-JUL-1996.  
 PE 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PR (AMGE-) AMGEN INC.  
 PA (AMGE-) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 DR WPI: 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 46-47; 79pp; English.  
 CC Deletives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. . They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 318 AA;

Query Match 100.0%; Score 40; DB 1; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 0.66;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 DB 53 VKALRVRL 61

RESULT 12  
 W00475  
 ID W00475 standard; protein; 152 AA.  
 AC W00475;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative (truncated).  
 KW adult respiratory distress syndrome; sepsis; liver failure;  
 KW CD14; treatment; prevention; inflammation; sepsis; acute pancreatitis;  
 KW adult respiratory distress syndrome; acute pancreatitis;

KW inflammatory bowel disease; graft versus host disease; bone marrow;  
 KW transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 PN W09620957-A1.  
 PD 11-JUL-1996.  
 PE 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PR (AMGE-) AMGEN INC.  
 PA (AMGE-) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 DR WPI: 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 47-48; 79pp; English.  
 CC Deletives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. . They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 152 AA;

Query Match 100.0%; Score 40; DB 1; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.32;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75

RESULT 13  
 W00476  
 ID W00476 standard; protein; 152 AA.  
 AC W00476;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative (truncated).  
 KW adult respiratory distress syndrome; sepsis; liver failure;  
 KW CD14; treatment; prevention; inflammation; sepsis; acute pancreatitis;  
 KW inflammatory bowel disease; graft versus host disease; bone marrow;  
 KW transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 PN W09620957-A1.  
 PD 11-JUL-1996.  
 PE 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PR (AMGE-) AMGEN INC.  
 PA (AMGE-) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 DR WPI: 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 48-49; 79pp; English.  
 CC Deletives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. . They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 152 AA;

PS Claim 1; Page 40-41; 79pp; English.  
 CC Derivatives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 348 AA;

Query Match 100.0%; Score 40; DB 1; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75

RESULT 8  
 W00471  
 ID W00471 standard; protein; 348 AA.  
 AC W00471;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative.  
 KM CD14; treatment: prevention; inflammation; sepsis; liver failure;  
 KM adult respiratory distress syndrome; acute pancreatitis;  
 KM inflammatory bowel disease; graft versus host disease; bone marrow;  
 KM transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.

Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 FT disulfide\_bond 168..198  
 FT disulfide\_bond 222..253  
 FT disulfide\_bond 287..333  
 PN W09620957-A1.  
 PD 11-JUL-1996.  
 PF 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PA (AMGE-) AMGEN INC.  
 PA (UYRO ) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;

DR WPI: 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 41-43; 79pp; English.  
 CC Derivatives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 348 AA;

Query Match 100.0%; Score 40; DB 1; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75

RESULT 9  
 W00472  
 ID W00472 standard; protein; 348 AA.  
 AC W00472;  
 DT 04-FEB-1997 (first entry)

DE Soluble CD14 derivative.  
 KM CD14; treatment: prevention; inflammation; sepsis; liver failure;  
 KM adult respiratory distress syndrome; acute pancreatitis;  
 KM inflammatory bowel disease; graft versus host disease; bone marrow;  
 KM transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.

Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 FT disulfide\_bond 168..198  
 FT disulfide\_bond 222..253  
 FT disulfide\_bond 287..333  
 PN W09620957-A1.  
 PD 11-JUL-1996.  
 PF 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PA (AMGE-) AMGEN INC.  
 PA (UYRO ) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;

DR WPI: 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 43-44; 79pp; English.  
 CC Derivatives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SQ Sequence 348 AA;

Query Match 100.0%; Score 40; DB 1; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75

RESULT 10  
 W00473  
 ID W00473 standard; protein; 348 AA.  
 AC W00473;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative.  
 KM CD14; treatment: prevention; inflammation; sepsis; liver failure;  
 KM adult respiratory distress syndrome; acute pancreatitis;  
 KM inflammatory bowel disease; graft versus host disease; bone marrow;  
 KM transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.

PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Figure 1; 79pp; English.  
 CC Derivatives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SO Sequence 348 AA;

Query Match 100.0%; Score 40; DB 1; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75

RESULT 5  
 ID W00468 standard; protein; 348 AA.  
 AC W00468;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative.  
 CC CD14; treatment; prevention; inflammation; sepsis; liver failure;  
 CC adult respiratory distress syndrome; acute pancreatitis;  
 CC inflammatory bowel disease; graft versus host disease; bone marrow;  
 CC transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 FT disulfide\_bond 168..198  
 FT disulfide\_bond 222..253  
 FT disulfide\_bond 287..333  
 PN W09620957-A1.  
 PD 11-JUL-1996.  
 PF 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PA (AMGE-) AMGEN INC.  
 PA (UYRO ) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 DR WPI; 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 37-38; 79pp; English.  
 CC Derivatives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SO Sequence 348 AA;

Query Match 100.0%; Score 40; DB 1; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75

RESULT 6  
 ID W00469 standard; protein; 348 AA.

AC W00469;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative.  
 CC CD14; treatment; prevention; inflammation; sepsis; liver failure;  
 CC adult respiratory distress syndrome; acute pancreatitis;  
 CC inflammatory bowel disease; graft versus host disease; bone marrow;  
 CC transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 FT disulfide\_bond 168..198  
 FT disulfide\_bond 222..253  
 FT disulfide\_bond 287..333  
 PN W09620957-A1.  
 PD 11-JUL-1996.  
 PF 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PA (AMGE-) AMGEN INC.  
 PA (UYRO ) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 DR WPI; 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis  
 PS Claim 1; Page 38-40; 79pp; English.  
 CC Derivatives of the CD14 polypeptide may be used to treat or prevent  
 CC inflammation, especially sepsis but also adult respiratory distress  
 CC syndrome, acute pancreatitis, acute or chronic liver failure,  
 CC inflammatory bowel disease, graft vs. host disease in bone marrow  
 CC transplant, tuberculosis etc. They may also be used for in vitro  
 CC detoxification and prevention of endotoxin contamination. See W00468-  
 CC 81.  
 SO Sequence 348 AA;

Query Match 100.0%; Score 40; DB 1; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75

RESULT 7  
 ID W00470 standard; protein; 348 AA.  
 AC W00470;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14 derivative.  
 CC CD14; treatment; prevention; inflammation; sepsis; liver failure;  
 CC adult respiratory distress syndrome; acute pancreatitis;  
 CC inflammatory bowel disease; graft versus host disease; bone marrow;  
 CC transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT disulfide\_bond 15..32  
 FT disulfide\_bond 168..198  
 FT disulfide\_bond 222..253  
 FT disulfide\_bond 287..333  
 PN W09620957-A1.  
 PD 11-JUL-1996.  
 PF 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PA (AMGE-) AMGEN INC.  
 PA (UYRO ) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 DR WPI; 96-333944/33.  
 PT Modified forms of CD14 polypeptide and DNA encoding them - useful  
 PT for treating and preventing inflammation, esp. sepsis

FT modified\_site /note= "potential N-linked glycosylation site"  
 FT 323..325  
 FT /label= Glycosylation  
 FT /note= "potential N-linked glycosylation site"  
 PN US5543303-A.  
 PD 06-AUG-1996.  
 PF 28-NOV-1988; US-276794.  
 PR 28-NOV-1988; US-276794.  
 PR 08-JUN-1990; US-536163.  
 PR 06-NOV-1991; US-787763.  
 PR 22-JUL-1992; US-916806.  
 PR 13-DEC-1993; US-165583.  
 PA (GOYER/) GOYERT S M.  
 PI Goyert SM;  
 DR WPI: 96-370638/37.  
 DR N-PSDB: T39716;  
 DR N-PSDB: T39717.  
 PT Recombinant DNA encoding myelomonocytic differentiation antigen CD14  
 PT - used for producing recombinant CD14 for use in e.g. diagnosis of  
 PT myeloid disorders such as leukaemia  
 PS Claim 1: Fig 3: 11pp; English.  
 CC Human myelomonocytic differentiation antigen CD14 (W05316) is an  
 CC antigen useful in the diagnosis of mature myeloid leukemia. Its  
 CC amino acid sequence was deduced from a cDNA clone (T39717) obt'd.  
 CC by screening COS 7 cell transfectants with monoclonal antibodies  
 CC to CD14. Large amts. of CD14 can be produced by expression in  
 CC transformed host cells; mature, glycosylated CD14 is produced in  
 CC mammalian host cells, and nonglycosylated CD14 in prokaryotic  
 CC hosts.  
 SQ Sequence 375 AA;

Query Match 100.0%; Score 40; DB 1; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 0.78;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
 DB 86 VKALRVRL 94  
 |||||

RESULT 3  
 ID R98576 standard; peptide; 348 AA.  
 AC R98576;  
 DT 07-FEB-1997 (first entry)  
 DE CD14 sequence.  
 KW Antiinflammatory; lipopolysaccharide; LPS; Gram-positive bacteria;  
 KW CD14; treatment; septic shock; inflammatory bowel disease;  
 KW liver failure; graft versus host disease; pancreatitis; tuberculosis;  
 KW adult respiratory distress syndrome; detection; quantification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Domain 1..56  
 FT /label= EGF-like domain  
 FT disulfide\_bond 6..15  
 FT disulfide\_bond 15..17  
 FT modified\_site 18  
 FT /note= "Potential N-linked glycosylation site"  
 FT disulfide\_bond 17..32  
 FT region 57..64  
 FT /label= Linker  
 FT region 65..283  
 FT /label= Core region  
 FT disulfide\_bond 168..198  
 FT disulfide\_bond 222..253  
 FT modified\_site 263  
 FT /note= "Potential N-linked glycosylation site"  
 FT disulfide\_bond 287..333  
 FT region 284..348  
 FT /label= tail  
 FT modified\_site 304  
 FT /note= "Potential N-linked glycosylation site"

PN W09620956-Al.  
 PD 11-JUL-1996.  
 PF 28-DEC-1995; U16977.  
 PR 30-DEC-1994; US-366953.  
 PR (AMGE-) AMGEN INC.  
 PA (UYRO ) UNIV ROCKEFELLER.  
 PI Juan S, Lichenstein HS, Narhi LO, Wright SD;  
 DR WPI: 96-333943/33.  
 PT Peptide(s) based on CD14 sequences - which bind to  
 PT lipo:poly:saccharide and inhibit CD14 mediated inflammatory  
 PT responses  
 PS Disclosure; Figure 1; 103pp; English.  
 CC Antinflammatory peptides based on the human CD14 sequence (See  
 CC R98570-75 and R98577-80) can bind to lipopolysaccharide (LPS) and  
 CC inhibit binding of LPS or Gram positive cell components to CD14.  
 CC thus reducing or eliminating CD14 mediated inflammatory responses.  
 CC They can be used for treating inflammatory conditions in particular,  
 CC septic shock, inflammatory bowel disease, acute and chronic liver  
 CC failure, graft versus host disease, intestinal or liver transplant,  
 CC adult respiratory distress syndrome, acute pancreatitis and  
 CC tuberculosis. They can also be used to remove, detect or quantitate  
 CC LPS or Gram-positive cell components in samples. The peptides are  
 CC used in doses of 0.1-100 mg/kg by parenteral or oral routes.  
 SQ Sequence 348 AA;

Query Match 100.0%; Score 40; DB 1; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKALRVRL 9  
 DB 67 VKALRVRL 75  
 |||||

RESULT 4  
 W00467  
 ID W00467 standard; protein; 348 AA.  
 AC W00467;  
 DT 04-FEB-1997 (first entry)  
 DE Soluble CD14.  
 KW CD14; treatment; prevention; inflammation; sepsis; liver failure;  
 KW adult respiratory distress syndrome; acute pancreatitis;  
 KW inflammatory bowel disease; graft versus host disease; bone marrow;  
 KW transplant; tuberculosis; endotoxin contamination; detoxification.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 6..17  
 FT misc\_difference 7  
 FT /note= "Gly or an amino acid with a C1-C6 aliphatic  
 FT side chain"  
 FT misc\_difference 8  
 FT /note= "Gly or an amino acid with a C1-C6 aliphatic  
 FT side chain"  
 FT misc\_difference 9  
 FT /note= "Gly or an amino acid with a C1-C6 aliphatic  
 FT side chain"  
 FT misc\_difference 10  
 FT /note= "Gly or an amino acid with a C1-C6 aliphatic  
 FT side chain"  
 FT disulfide\_bond 15..32  
 FT disulfide\_bond 168..198  
 FT disulfide\_bond 222..253  
 FT disulfide\_bond 287..333  
 PN W09620957-Al.  
 PD 11-JUL-1996.  
 PF 28-DEC-1995; U17095.  
 PR 30-DEC-1994; US-366953.  
 PR 07-JUN-1995; US-484397.  
 PA (AMGE-) AMGEN INC.  
 PA (UYRO ) UNIV ROCKEFELLER.  
 PI Juan SC, Lichenstein HS, Wright SD;  
 DR WPI: 96-333944/33.

Result No.	Score	Query		ID	Description
		Match	Length DB		
1	40	100.0	9	R71792	Peptide neutralist
2	40	100.0	375	W05316	Myelomonocytic dif
3	40	100.0	348	R98576	CD14 sequence. Pep
4	40	100.0	348	W00467	Soluble CD14. Modi
5	40	100.0	348	W00468	Soluble CD14 deriv
6	40	100.0	348	W00469	Soluble CD14 deriv
7	40	100.0	348	W00470	Soluble CD14 deriv
8	40	100.0	348	W00471	Soluble CD14 deriv
9	40	100.0	348	W00472	Soluble CD14 deriv
10	40	100.0	348	W00473	Soluble CD14 deriv
11	40	100.0	318	W00474	Soluble CD14 deriv
12	40	100.0	152	W00475	Soluble CD14 deriv
13	40	100.0	152	W00476	Soluble CD14 deriv
14	40	100.0	152	W00477	Soluble CD14 deriv
15	40	100.0	152	W00478	Soluble CD14 deriv
16	40	100.0	152	W00479	Soluble CD14 deriv
17	40	100.0	152	W00480	Soluble CD14 deriv
18	40	100.0	138	W00481	Soluble CD14 deriv
19	40	100.0	9	W21609	Antibiotic potentl
20	40	100.0	375	W41693	Human CD14. Assay
21	40	100.0	375	W60854	Human CD14 protein
22	39	97.5	373	W60853	Bovine CD14 protei
23	33	82.5	462	R12401	Enantioselective a
24	33	82.5	462	R25320	Enantioselective a
25	32	80.0	1407	R63631	Murine receptor-ty
26	32	80.0	1457	R63633	Murine receptor-ty
27	32	80.0	1439	R63632	Human receptor-ty
28	30	75.0	2942	W22050	Saccharomyces cere
29	30	75.0	403	W72844	Polymerase enhanci
30	29	72.5	286	R85441	Neurotrophin-6. DN
31	29	72.5	479	W52863	Glycosaminoglycan
32	27	67.5	648	R22561	Mouse mutant c-ra
33	27	67.5	648	R22563	Mouse c-raf-1. Det
34	27	67.5	648	R22559	Human c-raf-1. Det
35	27	67.5	648	R25277	Mouse mutant c-ra
36	27	67.5	648	R22560	Mouse mutant c-ra
37	27	67.5	648	R38154	Acetobacter diquan
38	27	67.5	574	R22562	Seven transmembran
39	27	67.5	339	R53752	Alginic acid lyase
40	27	67.5	622	R67434	Raf(51-131) oncopro
41	27	67.5	81	R66656	Raf(1-257) oncopro
42	27	67.5	257	R56644	Raf(1-149) oncopro
43	27	67.5	149	R56645	Raf(1-149) oncopro

## LE



**THIS PAGE BLANK (USPTO)**

us-09-124-280a-20.rspt

Wed Sep 8 09:43:02 1999

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
 OC Drosophilidae; Drosophila.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP YOSHIDA E.N., BENKEL B.F., FONG Y., HICKEY D.A.;  
 RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF094764; AAC95306.1;  
 KW Kinase.  
 SQ SEQUENCE 647 AA; 71592 MW; 24AF8DE1 CRC32;

Query Match 54.4%; Score 37; DB 5; Length 647;  
 Best Local Similarity 60.0%; Pred. No. 94;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 FLKFLKFLKC 12  
 DB 171 FVKFFRFHKC 180

RESULT 14  
 Q98140 PRELIMINARY; PRT; 752 AA.  
 AC Q98140;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE GENE 24 PROTEIN.  
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae; Rhadinovirus.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 96099469.  
 RX MOORE P.S., GAO S.J., DOMINGUEZ G., CESARMAN E., LUNGU O.,  
 RA KNOWLES D.M., GARBER R., PELLET P.E., MCGEOCH D.J., CHANG Y.;  
 RT "Primary characterization of a herpesvirus agent associated with  
 Kaposi's sarcoma."  
 RL J. Virol. 70:549-558(1996).  
 DR EMBL: U40377; AAB08393.1;  
 SQ SEQUENCE 752 AA; 85211 MW; E53834A0 CRC32;

Query Match 54.4%; Score 37; DB 12; Length 752;  
 Best Local Similarity 41.7%; Pred. No. 1.le+02;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 CKFLKFLKFLKC 12  
 DB 165 CRYLEFWEFEQC 176

RESULT 15  
 P88916 PRELIMINARY; PRT; 752 AA.  
 AC P88916;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE ORF 24.  
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae; Rhadinovirus.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 96099469.  
 RX MOORE P.S., GAO S.J., DOMINGUEZ G., CESARMAN E., LUNGU O.,  
 RA KNOWLES D.M., GARBER R., PELLET P.E., MCGEOCH D.J., CHANG Y.;  
 RT "Primary characterization of a herpesvirus agent associated with  
 Kaposi's sarcoma."  
 RL J. Virol. 70:549-558(1996).

[2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 97111480.  
 RX RUSSO J.J., BOHENZKY R.A., CHIEN M.C., CHEN J., YAN M., MADDALENA D.,  
 RA PARRY J.P., PERUZZI D., EDELMAN I.S., CHANG Y., MOORE P.S.;  
 RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus  
 (HHV8)."  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 97094384.  
 RX MOORE P.S., BASHOFF G., WEISS R.A., CHANG Y.;  
 RT "Molecular mimicry of human cytokine and cytokine response pathway  
 genes by KSHV."  
 RL Science 274:1739-1744(1996).  
 DR EMBL: U75698; AAC57105.1;  
 SQ SEQUENCE 752 AA; 85192 MW; DCD6FC4F CRC32;

Query Match 54.4%; Score 37; DB 12; Length 752;  
 Best Local Similarity 41.7%; Pred. No. 1.le+02;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 CKFLKFLKFLKC 12  
 DB 165 CRYLEFWEFEQC 176

Search completed: September 7, 1999, 20:34:53  
 Job time: 19749 sec

OC Viruses: dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;  
 OC Entomopoxvirus B.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=TUCSON;  
 RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.;  
 RT "The genome of Melanoplus sanguinipes entomopoxvirus.";  
 RL J. Virol. 73:533-552(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=TUCSON;  
 RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF063866; AAC97718.1; -.  
 SQ SEQUENCE 429 AA; 50651 MW; 216304FE CRC32;

Query Match 55.9%; Score 38; DB 12; Length 429;  
 Best Local Similarity 72.7%; Pred. No. 47;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CKFLKFLKFLK 11  
 ||||| || |:  
 Db 69 CKFLKSLKLE 79

RESULT 11  
 ID O07568 PRELIMINARY; PRT; 384 AA.  
 AC O07568;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 41.0 KD PROTEIN.  
 GN YHJN.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/staphylococcus group; Bacillus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA NOBACK M.A., TERPSTRA P., HOLSAPPEL S., VENEMA G., BRON S.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE: 98044033.  
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
 RA BROUILLET S., BRUSCH C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., ENMERSON P.T.,  
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,  
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,  
 RA GHM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
 RA GUISEPI G., GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A.,  
 RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,  
 RA JORIS B., KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C.,  
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
 RA NOONE D., O'REILLY M., OGAWA K., OGIWARA A., OUDEGA B., PARK S.H.,  
 RA PARRO V., POHL T.M., PORTETELLE D., PORWOLLIK S., PRESCOTT A.M.,  
 RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
 RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,  
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,  
 RA SERIGUCHI J., SEKOWSKA A., SERO S.J., SERROR P., SHIN B.S., SOLDI B.,  
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
 RA TAKEUCHI M., TANAKASHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
 RA TORATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
 RA VIARI A., WAMBUTT R., WEDLER E., WEDLER H., WEITZENEGGER T.,

RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,  
 RA YOSHIDA K., YOSHIKAWA H.F., ZUNSTEIN E., YOSHIKAWA H., DANCHIN A.;  
 RT "The complete genome sequence of the gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL Nature 390:249-256(1997).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=168;  
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Y14081; CAA74476.1; -.  
 DR EMBL: Z99109; CAB12897.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 384 AA; 41029 MW; FC2C3080 CRC32;

Query Match 54.4%; Score 37; DB 2; Length 384;  
 Best Local Similarity 77.8%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CKFLKFLKFLK 9  
 || |||||  
 Db 214 CKAFLKFLK 222

RESULT 12  
 ID O51627 PRELIMINARY; PRT; 360 AA.  
 AC O51627;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE CAROTENOID BIOSYNTHESIS PROTEIN, PUTATIVE.  
 GN BB0684.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE: 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia  
 RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 RL EMBL: AE001169; AAC67033.1; -.  
 DR TIGR: BB0684; -.  
 SQ SEQUENCE 360 AA; 40013 MW; 7CD8AAE2 CRC32;

Query Match 54.4%; Score 37; DB 2; Length 360;  
 Best Local Similarity 75.0%; Pred. No. 58;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CKFLKFLK 8  
 | |||||  
 Db 34 CNFLKFLK 41

RESULT 13  
 ID O96613 PRELIMINARY; PRT; 647 AA.  
 AC O96613;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE SNF4/AMP-ACTIVATED PROTEIN KINASE GAMMA SUBUNIT.

01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
SIMILAR TO AMSACTA

GN PH0064.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OT3;  
 RX MEDLINE; 98344137.  
 RA KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
 RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,  
 RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
 RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OUCHI A.,  
 RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
 RA KIKUCHI H.;  
 RT \*Complete sequence and gene organization of the genome of a  
 RT hyperthermophilic archaeobacterium, Pyrococcus horikoshii OT3.\*;  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL; AP000001; BAA29133.1; -.  
 SQ SEQUENCE .102 AA; 11083 MW; AB79A987 CRC32;

Query Match 58.8%; Score 40; DB 1; Length 102;  
 Best Local Similarity 80.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 FLKFLKFLK 12  
 Db 38 FLTLKSLK 47  
 |||||  
 |||||

RESULT 3  
 ID O96277 PRELIMINARY; PRT; 657 AA.  
 AC O96277;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE PROTEIN WITH DNAJ DOMAIN (RESA-LIKE).  
 GN PF00925W.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
 RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
 RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
 RT falciparum";  
 RL Science 282:1126-1132(1998).  
 DR EMBL; AE001426; AAC71974.1; -.  
 SQ SEQUENCE 657 AA; 77802 MW; 75B5E5A3 CRC32;

Query Match 58.8%; Score 40; DB 5; Length 657;  
 Best Local Similarity 60.0%; Pred. No. 32;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKFLKFLKFL 10  
 Db 642 CRLKFLKYM 651  
 |:|||||:

RESULT 4  
 ID O66417 PRELIMINARY; PRT; 193 AA.  
 AC O66417;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 22.6 KD PROTEIN.  
 GN AA26.  
 OS Aquifex aeolicus.

OG Plasmid ecel.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus";  
 RL Nature 392:353-358(1998).  
 DR EMBL; AE000667; AAC07969.1; -.  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 193 AA; 22552 MW; 935BAF96 CRC32;

Query Match 57.4%; Score 39; DB 2; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 FLKFLK 12  
 Db 57 FLKFLK 63  
 |||||  
 |||||

RESULT 5  
 ID O44194 PRELIMINARY; PRT; 261 AA.  
 AC O44194;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE T11F8.1 PROTEIN.  
 GN T11F8.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718;  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA JONES M., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SALLDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA JOHNSON D.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF036703; AAB88555.1; -.  
 SQ SEQUENCE 261 AA; 30914 MW; 7DBE6570 CRC32;

Query Match 57.4%; Score 39; DB 5; Length 261;  
 Best Local Similarity 60.0%; Pred. No. 22;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 FLKFLKFLK 12

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:52 ; Search time 148.39 Seconds  
(without alignments)  
4.977 Million cell updates/sec

Title: US-09-124-280A-20  
Perfect score: 68  
Sequence: 1 CKFLKFLKFLKC 12

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL10:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	60.3	240	1 Q58624	Q58624 methanococ
2	40	58.8	102	1 O57800	O57800 pyrococcus
3	40	58.8	657	5 O96277	O96277 plasmodium
4	39	57.4	193	2 O66417	O66417 aquifex ae
5	39	57.4	261	5 O44194	O44194 caenorhabdi
6	39	57.4	1531	10 O04565	O04565 arabidopsis
7	39	57.4	1873	12 Q83044	Q83044 lettuce inf
8	38	55.9	59	2 Q46199	Q46199 clostridium
9	38	55.9	1944	3 O94387	O94387 schizosacch
10	38	55.9	429	12 Q9YVK3	Q9YVK3 melanoplus
11	37	54.4	384	2 O07568	O07568 bacillus su
12	37	54.4	360	2 O51627	O51627 borrelia bu
13	37	54.4	647	5 O96613	O96613 drosophila
14	37	54.4	752	12 Q28140	Q28140 kaposi's sa
15	37	54.4	752	12 P88916	P88916 kaposi's sa
16	37	54.4	752	12 O40924	O40924 kaposi's sa
17	36.5	53.7	264	5 Q20246	Q20246 caenorhabdi
18	36	52.9	150	2 O51471	O51471 borrelia bu
19	36	52.9	902	3 Q12482	Q12482 saccharomyc
20	36	52.9	181	5 O76674	O76674 caenorhabdi
21	36	52.9	359	5 Q20831	Q20831 caenorhabdi
22	36	52.9	110	6 P79157	P79157 callitrix
23	36	52.9	464	6 O46520	O46520 sus scrofa
24	36	52.9	951	10 Q43106	Q43106 phaseolus v
25	36	52.9	953	10 P93265	P93265 mesembryant
26	36	52.9	956	10 Q42932	Q42932 nicotiana p
27	36	52.9	963	10 Q43131	Q43131 vicia faba
28	36	52.9	956	10 Q43182	Q43182 solanum tub
29	36	52.9	949	10 Q43243	Q43243 zea mays (m

30	36	52.9	956	10 Q43001	Q43001 oryza sativ
31	36	52.9	957	10 Q43002	Q43002 oryza sativ
32	36	52.9	472	11 O61066	O61066 mus musculu
33	36	52.9	403	13 O73742	O73742 xenopus lae
34	35.5	52.2	2437	5 O77393	O77393 plasmodium
35	35	51.5	1279	2 O26046	O26046 helicobacte
36	35	51.5	416	3 O59832	O59832 schizosacch
37	35	51.5	290	3 P78815	P78815 schizosacch
38	35	51.5	820	4 O60585	O60585 homo sapien
39	35	51.5	327	5 O22984	O22984 caenorhabdi
40	35	51.5	546	5 O45887	O45887 caenorhabdi
41	35	51.5	410	8 O21257	O21257 reclinomona
42	35	51.5	552	8 Q92YM7	Q92YM7 rhipicephal
43	35	51.5	1299	10 P93098	P93098 ceratodon p
44	35	51.5	314	10 O82023	O82023 medicago tr
45	35	51.5	897	11 O70495	O70495 mus musculu

ALIGNMENTS

RESULT 1

Q58624 Q58624 PRELIMINARY; PRT; 240 AA.  
ID Q58624;  
AC Q58624;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)  
DE HYPOTHETICAL PROTEIN M1227.  
GN M1227.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96337999.

RA BULT C.-J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
RA SCOTT J.L., GEORGE N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,  
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
RA \*Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii.\*  
RT Science 273:1058-1073(1996).  
RL -1- SIMILARITY: TO THE ORGANIC RADICAL ACTIVATING ENZYMES FAMILY.  
CC EMBL: U67563; AAB99230.1; -  
DR Hypothetical protein; Iron.  
KW Hypothetical protein; Iron.  
FT METAL 29 29 IRON (POTENTIAL).  
FT METAL 33 33 IRON (POTENTIAL).  
FT METAL 36 36 IRON (POTENTIAL).  
SQ SEQUENCE 240 AA; 27682 MW; 42EE7760 CRC32;

Query Match 60.3%; Score 41; DB 1; Length 240;

Best Local Similarity 50.0%; Pred. No. 9.7;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKFLKFLKC 12

Db 124 CRFDKYEFVKC 135

RESULT 2

O57800 O57800 PRELIMINARY; PRT; 102 AA.

ID O57800;  
AC O57800;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
DE 102AA LONG HYPOTHETICAL PROTEIN.

QY 2 KFLKFLKFL 10  
||||| I:  
Db 61 KFLKFLGFM 69

## RESULT 12

S50751  
H+-transporting ATPase (EC 3.6.1.35) (clone PHA1) - potato  
C:Species: Solanum tuberosum (potato)  
C:Date: 14-Jul-1995 #sequence\_revision 08-Sep-1995 #text\_change 24-Apr-1998  
C:Accession: S50751; S39313  
R:Harms, K.; Woelner, R.V.; Schulz, B.; Frommer, W.B.  
Plant Mol. Biol. 26, 979-988, 1994  
A:Title: Isolation and characterization of P-type H(+)-ATPase genes from potato.  
A:Reference number: S50751  
A:Accession: S50751  
A:Molecule type: mRNA  
A:Residues: 1-956 <HAR>  
A:Cross-references: EMBL:X76536; NID:g435002; PID:g435003  
A:Note: the nucleotide is not given  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; hydrolase; phosphoprotein  
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:427/Binding site: ATP (Lys) #status predicted

Query Match 52.9%; Score 36; DB 2; Length 956;  
Best Local Similarity 77.8%; Pred. No. 96;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
||||| I:  
Db 61 KFLKFLGFM 69

## RESULT 13

S52728  
H+-transporting ATPase (EC 3.6.1.35) - kidney bean  
C:Species: Phaseolus vulgaris (kidney bean)  
C:Date: 19-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 24-Apr-1998  
C:Accession: S52728  
R:Campos, F.; Perez-Castineira, J.R.; Villalba, J.M.; Serrano, R.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S52728  
A:Accession: S52728  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-951 <CAM>  
A:Cross-references: EMBL:X85804; NID:g758249; PID:g758250  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; hydrolase; phosphoprotein  
F:483-655/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:331/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:425/Binding site: ATP (Lys) #status predicted

Query Match 52.9%; Score 36; DB 2; Length 951;  
Best Local Similarity 77.8%; Pred. No. 95;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
||||| I:  
Db 59 KFLKFLGFM 67

## RESULT 14

S54495  
probable carrier protein YPR021c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YP9367.01c  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Apr-1998  
C:Accession: S54495

R:Badcock, K.; Churcher, C.M.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54059  
A:Accession: S54495  
A:Molecule type: DNA  
A:Residues: 1-902 <BAD>  
A:Cross-references: EMBL:Z49274; NID:g809585; PID:g809586; MIPS:YPR021c  
A:Experimental source: strain AB972  
C:Genetics:  
A:Map position: 16R  
C:Superfamily: probable carrier protein YPR021c; ADP,ATP carrier protein repeat homol  
C:Keywords: duplication; transmembrane protein  
F:527-615/Domain: ADP,ATP carrier protein repeat homology <ACP1>  
F:621-711/Domain: ADP,ATP carrier protein repeat homology <ACP2>  
F:724-814/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 52.9%; Score 36; DB 2; Length 902;  
Best Local Similarity 75.0%; Pred. No. 91;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 KFLKFLKC 12  
||||| I:  
Db 223 RFLKFLKFC 230

## RESULT 15

H70164  
hypothetical protein BB0521 - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
C:Accession: H70164  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: H70164  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-150 <KLE>  
A:Cross-references: GB:AE001154; GB:AE000783; NID:g2688431; PID:g2688446; TIGR:BB0521  
A:Experimental source: strain B31

Query Match 52.9%; Score 36; DB 2; Length 150;  
Best Local Similarity 62.5%; Pred. No. 18;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKFLK 8  
||||| I:  
Db 63 CKFIEPMK 70

Search completed: September 7, 1999, 23:06:34  
Job time: 2482 sec

A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.

A:Reference number: A70100; MUID:98065943

A:Accession: C70185

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-360 <KLE>

A:Cross-references: GB:AE001169; GB:AE000783; NID:g26886117; PID:g26886117; TIGR:BB0684

A:Experimental source: strain B31

C:Superfamily: carotenoid biosynthesis protein homolog

Query Match 54.4%; Score 37; DB 2; Length 360;  
Best Local Similarity 75.0%; Pred. No. 27;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFLKFLK 8

Db 34 CNFLKFLK 41

RESULT 8

E59834 hypothetical protein yhjN - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998

C:Accession: E59834

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

iochi, M.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamamoto, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; MUID:98044033

A:Accession: E59834

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-384 <KUN>

A:Cross-references: GB:Z99109; GB:AL009126; NID:g26333260; PID:e1183059; PID:g26333393

C:Experimental source: strain 168

C:Genetics:

A:Gene: yhjN

Query Match 54.4%; Score 37; DB 2; Length 384;  
Best Local Similarity 77.8%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKFLK 9

Db 214 CKAFLKLF 222

RESULT 9

B39403 probable homocitrate synthase (EC 4.1.3.21) Nifv omega chain - *Clostridium pasteurianum*

C:Species: *Clostridium pasteurianum*

C:Date: 28-Feb-1992 #sequence\_revision 28-Feb-1992 #text\_change 20-Mar-1998

C:Accession: B39403

R:Wang, S.Z.; Dean, D.R.; Chen, J.S.; Johnson, J.L.

J. Bacteriol. 173, 3041-3046, 1991

A:Title: the N-terminal and C-terminal portions of Nifv are encoded by two different gen

A:Reference number: A39403; MUID:91216974

A:Accession: B39403

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-269 <WAN>

A:Cross-references: GB:M68878; NID:g144877; PID:g144879

C:Keywords: carbon-carbon lyase; oxo-acid-lyase

Query Match 54.4%; Score 37; DB 2; Length 269;  
Best Local Similarity 54.5%; Pred. No. 21;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKFLK 11

Db 229 CDFLKFKVQ 239

RESULT 10

A45506 H+-transporting ATPase (EC 3.6.1.35) LHA1 - tomato

C:Species: *Lycopersicon esculentum* (tomato)

C:Date: 14-May-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Mar-1998

C:Accession: A45506

R:Swing, N.N.; Wimmers, L.E.; Meyer, D.J.; Chetelat, R.T.; Bennett, A.B.

Plant Physiol. 94, 1874-1881, 1990

A:Title: Molecular cloning of tomato plasma membrane H+-ATPase.

A:Reference number: A45506

A:Accession: A45506

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-956 <EWI>

A:Cross-references: GB:M60166; NID:g170463; PID:g170464

C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding doma

C:Keywords: ATP; hydrolase; phosphoprotein; transmembrane protein

F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>

F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:427/Binding site: ATP (Lys) #status predicted

Query Match 52.9%; Score 36; DB 2; Length 956;  
Best Local Similarity 77.8%; Pred. No. 96;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLK 10

Db 61 KFLKFLGFM 69

RESULT 11

A43637 H+-transporting ATPase (EC 3.6.1.35) - curled-leaved tobacco

C:Species: *Nicotiana glauca* (curled-leaved tobacco)

C:Date: 11-Feb-1993 #sequence\_revision 12-Mar-1993 #text\_change 17-Mar-1999

C:Accession: A43637

R:Boutry, M.; Michelet, B.; Goffeau, A.

Biochem. Biophys. Res. Commun. 162, 567-574, 1989

A:Title: Molecular cloning of a family of plant genes encoding a protein homologous t

A:Reference number: A43637; MUID:89334860

A:Accession: A43637

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-956 <BOU>

A:Cross-references: GB:M27888; NID:g170205; PID:g170206

A:Note: the authors translated the codon CGG for residue 131 as Trp

C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding doma

C:Keywords: ATP; hydrolase; phosphoprotein

F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>

F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:427/Binding site: ATP (Lys) #status predicted

Query Match 52.9%; Score 36; DB 2; Length 956;  
Best Local Similarity 77.8%; Pred. No. 96;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;



Db 237 CKFLSFIRF 245  
|||||

## RESULT 3

G71602  
protein with DnaJ domain (RESA-like) PFB0925w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 16-Dec-1998  
C:Accession: G71602  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;  
Pierce, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600.  
A:Accession: G71602  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-657 <GAR>  
A:Cross-references: GB:AF001426; GB:AE001362; NID:g3845312; PID:g3845313; TIGR:PFB0925w  
A:Experimental source: clone 307  
C:Genetics:  
A:Gene: PFB0925w  
C:Superfamily: dnaJ amino-terminal homology  
F:246-311/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 58.8%; Score 40; DB 2; Length 657;

Best Local Similarity 60.0%; Pred. No. 15;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKFLKFLKFL 10

Db 642 CKLKLKFLKYM 651

|||||

## RESULT 4

F71225  
hypothetical protein PH0064 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 14-Aug-1998  
C:Accession: F71225  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137  
A:Accession: F71225  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-102 <KAW>  
A:Cross-references: GB:AF000001; NID:g3236128; PID:d1030076; PID:g3256450  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0064

Query Match 58.8%; Score 40; DB 2; Length 102;

Best Local Similarity 80.0%; Pred. No. 2.7;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 FLKFLKFLKLC 12

Db 38 FLTFLKSLKLC 47

|||||

## RESULT 5

F69888  
alanine racemase homolog yncD - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998  
C:Accession: F69888

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal  
lech, J.; Harwood, C.R.; Menaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maguda, S.; Mau  
Y, M.; Ogawa, K.; Ogilwa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yaguchi, K.; Yata, K.; Yoshiya  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
A:Reference number: A69580; MUID:98044033  
A:Accession: F69888  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-394 <KUN>  
A:Cross-references: GB:Z99113; GB:AL009126; NID:g2634090; PID:el183423; PID:g2634148  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yncD  
C:Superfamily: alanine racemase

Query Match 55.9%; Score 38; DB 2; Length 394;

Best Local Similarity 70.0%; Pred. No. 20;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11

Db 187 KFTSFLSFLK 196

|||||

## RESULT 6

I40860  
hypothetical protein - Clostridium pasteurianum (fragment)  
C:Species: Clostridium pasteurianum  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Aug-1996  
C:Accession: I40860  
R:Daldal, F.; Applebaum, J.  
J. Mol. Biol. 186, 533-545, 1985  
A:Title: Cloning and expression of Clostridium pasteurianum galactokinase gene in Esc  
A:Reference number: I40860; MUID:86143810  
A:Accession: I40860  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-59 <RES>  
A:Cross-references: EMBL:X03352; NID:g40566; PID:g40367

Query Match 55.9%; Score 38; DB 2; Length 59;

Best Local Similarity 58.3%; Pred. No. 3.6;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CKFLKFLKFLKLC 12

|||||

Db 28 CKFLYLFLFNKC 39

## RESULT 7

C70185  
Carotenoid biosynthesis protein homolog - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 21-Aug-1998  
C:Accession: C70185  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.

Wed\_Sep 8 09:43:00 1999

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:33 ; Search time 116.8 Seconds  
(without alignments)  
4.116 Million cell updates/sec

Title: US-09-124-280A-20  
Perfect score: 68  
Sequence: 1 CKFLKFLKFLKC 12

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	41	60.3	240	1 B64453
2	40	58.8	309	2 B38545
3	40	58.8	657	2 G71602
4	40	58.8	102	2 F71225
5	38	55.9	394	2 F69888
6	38	55.9	59	2 I40860
7	37	54.4	360	2 C70185
8	37	54.4	384	2 E69834
9	37	54.4	269	2 B39403
10	36	52.9	956	2 A45506
11	36	52.9	956	2 A43637
12	36	52.9	956	2 S50751
13	36	52.9	951	2 S52728
14	36	52.9	902	2 S54495
15	36	52.9	150	2 H70164
16	36	52.9	84	2 S71290
17	36	52.9	949	2 T02083
18	36	52.9	543	2 S58095
19	36	52.9	803	2 S45916
20	36	52.9	470	2 S50854
21	36	52.9	470	2 J01960
22	36	52.9	472	2 GC1599
23	36	52.9	809	2 S40460
24	36	52.9	682	2 S40459
25	35	51.5	405	1 W1WLB4
26	35	51.5	1303	2 S27396
27	35	51.5	579	2 S35237
28	35	51.5	1279	2 E64709
29	35	51.5	338	2 S45908
30	35	51.5	483	2 S45347
31	35	51.5	410	2 S78151
32	34.5	50.7	365	2 S68679
33	34.5	50.7	344	2 S50474
34	34	50.0	1135	1 JQ1928
35	34	50.0	233	2 A23729
36	34	50.0	260	2 A69068
37	34	50.0	201	2 G70130
38	34	50.0	714	2 S76082
39	34	50.0	203	2 S47139

40 34 50.0 452 2 T02492  
41 34 50.0 583 2 S43139  
42 34 50.0 549 2 S46029  
43 34 50.0 765 2 S38099  
44 34 50.0 309 2 S61665  
45 34 50.0 578 2 S55102

hypothetical prote  
phycoene dehydroge  
hypothetical prote  
hypothetical prote  
probable membrane  
PSP2 protein - yea

ALIGNMENTS

RESULT 1  
B64453  
probable pyruvate formate-lyase activating enzyme (EC 1.97.4.1) - Methanococcus janna  
N.Alternate names: formate C-acetyltransferase activating enzyme; pyruvate formate-ly  
C.Species: Methanococcus jannaschii  
C.Date: 15-Jan-1999 #sequence\_revision 15-Jan-1999 #text\_change 15-Jan-1999  
C.Accession: B64453  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273 1058-1073, 1996  
A.Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
A.Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc  
A.Reference number: A64300; MUID:96337999  
A.Accession: B64453  
A>Status: nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-240 <BUL>  
A.Cross-references: GB:U67563; GB:L77117; NID:G1591850; PID:G1591858; TIGR:MJ1227; PI  
C.Genetics:  
C.Map position: FOR1170387-1171109  
A.Start codon: GTG  
A.Superfamily: Methanococcus probable pyruvate formate-lyase activating enzyme  
C.Keywords: iron; metalloprotein; oxidoreductase  
E;29,33,36/Binding site: iron (Cys) #status predicted

Query Match 60.3%; Score 41; DB 1; Length 240;  
Best Local Similarity 50.0%; Pred. No. 4;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKFLKFLKC 12  
|:| |:|:|:|  
Db 124 CRFDKYKEFKVC 135

RESULT 2  
B38545  
hypothetical protein 2 - maize rough dwarf virus  
C.Species: maize rough dwarf virus  
C.Date: 27-Nov-1991 #sequence\_revision 27-Nov-1991 #text\_change 09-Sep-1997  
C.Accession: B38545  
R:Marzachi, C.; Boccardo, G.; Nuss, D.L.  
Virology 180, 518-526, 1991  
A.Title: Cloning of the maize rough dwarf virus genome: molecular confirmation of the  
e genomic segment.  
A.Reference number: A38545; MUID:91111970  
A.Accession: B38545  
A>Status: preliminary  
A.Molecule type: genomic RNA  
A.Residues: 1-309 <MAR>  
A.Cross-references: EMBL:X55701; NID:G58388; PID:G58390  
A.Note: This virus is a member of the Reovirus group  
C.Genetics:  
A.Map position: segment 6

Query Match 58.8%; Score 40; DB 2; Length 309;  
Best Local Similarity 66.7%; Pred. No. 7.4;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFLKFLKFLK 9

; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 752 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-363-327-21

Query Match 54.4%; Score 37; DB 22; Length 752;  
Best Local Similarity 41.7%; Pred. No. 3.6e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKFLKFLKFLKC 12  
|:|:|:|:|:|:  
Db 165 CRYLEFWEFQC 176

RESULT 15  
US-09-270-767-57424  
; Sequence 57424, Application US/09270767  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 57424  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-57424

Query Match 53.7%; Score 36.5; DB 22; Length 337;  
Best Local Similarity 61.5%; Pred. No. 1.9e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 CKFLKF-LKFLKC 12  
|:|:|:|:|:|:  
Db 235 CKWLSFQTKFLHC 247

Search completed: September 7, 1999, 20:31:35  
Job time: 19927 sec

us-09-124-280a-20.rap

Wed Sep 8 09:42:59 1999

```

; MOLECULE TYPE: protein
; US-08-793-624A-21

Query Match      54.4%; Score 37; DB 9; Length 752;
Best Local Similarity 41.7%; Pred. No. 3.6e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,963B
FILING DATE: January 29, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45185-E/JPW/MS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 752 amino acids
TYPE: amino acid
TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-592-963B-21

Query Match      54.4%; Score 37; DB 9; Length 752;
Best Local Similarity 41.7%; Pred. No. 3.6e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKELKFLKFLKC 12
Db 165 CRYLEWFEQEC 176

```

```

; MOLECULE TYPE: protein
; US-08-793-624A-21

Query Match      54.4%; Score 37; DB 22; Length 153;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 KFLKFLKC 12
Db 62 EFLRFLKC 69

RESULT 14
US-09-363-327-21
; Sequence 21, Application US/09363327
; GENERAL INFORMATION:
; APPLICANT: Yuan Chang and Patrick S. Moore
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/363,327
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,963
; FILING DATE: January 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-E/JPW/MS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400

```

```

; MOLECULE TYPE: protein
; US-08-793-624A-21

Query Match      54.4%; Score 37; DB 9; Length 752;
Best Local Similarity 41.7%; Pred. No. 3.6e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKELKFLKFLKC 12
Db 165 CRYLEWFEQEC 176

RESULT 12
US-08-793-624A-21
; Sequence 21, Application US/08793624A
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,624A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 752 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/151138  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 45185-D-PCT/JPW/MS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 752 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-15138-21

Query Match 54.4%; Score 37; DB 1; Length 752;  
Best Local Similarity 41.7%; Pred. No. 3.6e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKFLKFLKFLKC 12  
|:|:|:|:|:|  
Db 165 CRYLEFWEFEQC 176

RESULT 9  
US-08-420-235-21  
; Sequence 21, Application US/08420235  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 752 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-420-235-21

Query Match 54.4%; Score 37; DB 7; Length 752;  
Best Local Similarity 41.7%; Pred. No. 3.6e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CKFLKFLKFLKC 12  
|:|:|:|:|:|  
Db 165 CRYLEFWEFEQC 176  
RESULT 10  
US-08-420-235A-21  
; Sequence 21, Application US/08420235A  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Moore, Patrick S.  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 752 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-420-235A-21

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/420,235A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 45185-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 752 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-420-235A-21

Query Match 54.4%; Score 37; DB 7; Length 752;  
Best Local Similarity 41.7%; Pred. No. 3.6e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKFLKFLKFLKC 12  
|:|:|:|:|:|  
Db 165 CRYLEFWEFEQC 176

RESULT 11  
US-08-592-963B-21  
; Sequence 21, Application US/08592963B  
; GENERAL INFORMATION:  
; APPLICANT: Yuan Chang and Patrick S. Moore  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 752 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-592-963B-21

us-09-124-280a-20.rap

Wed Sep 8 09:42:59 1999

```

; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 16944
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (28),(34),(42)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are un
US-09-248-796-16944

Query Match 55.9%; Score 38; DB 16; Length 635;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11
Db 533 KYLKILKFLK 542

RESULT 7
US-60-096-409-16944
; Sequence 16944, Application US/60096409A
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI
; FILE REFERENCE: GTC-016P
; CURRENT APPLICATION NUMBER: US/60/096.409A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 16944
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (28),(34),(42)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are un
US-60-096-409-16944

Query Match 55.9%; Score 38; DB 18; Length 635;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11
Db 533 KYLKILKFLK 542

RESULT 8
PCT-US95-15138-21
; Sequence 21, Application PC/TUS9515138
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Columbia University in the City of New York
; APPLICANT: City
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

Best Local Similarity 100.0%; Pred. No. 7.3e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKFLKFLKC 12  
Db 1 CKFLKFLKFLKC 12

RESULT 2  
US-09-124-280A-8  
; Sequence 8, Application US/09124280A  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: VACCINES FOR PREVENTION OF GRAM-  
; TITLE OF INVENTION: NEGATIVE BACTERIAL INFECTIONS AND ENDOTOXIN RELATED DISEASES  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/124,280A  
; FILING DATE: July 29,1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-09-124-280A-8

Query Match 73.5%; Score 50; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.04;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11  
Db 1 KFLKFLKFLK 10

RESULT 3  
US-09-124-280A-37  
; Sequence 37, Application US/09124280A  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: VACCINES FOR PREVENTION OF GRAM-  
; TITLE OF INVENTION: NEGATIVE BACTERIAL INFECTIONS AND ENDOTOXIN RELATED DISEASES  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA

ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/124,280A  
; FILING DATE: July 29,1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-09-124-280A-37

Query Match 73.5%; Score 50; DB 15; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.044;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11  
Db 2 KFLKFLKFLK 11

RESULT 4  
US-09-124-280A-41  
; Sequence 41, Application US/09124280A  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: VACCINES FOR PREVENTION OF GRAM-  
; TITLE OF INVENTION: NEGATIVE BACTERIAL INFECTIONS AND ENDOTOXIN RELATED DISEASE  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/124,280A  
; FILING DATE: July 29,1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 41:  
; SEQUENCE CHARACTERISTICS:

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:31:34 ; Search time 283.23 Seconds  
(without alignments)  
2.386 Million cell updates/sec

Title: US-09-124-280A-20  
Perfect score: 68  
Sequence: 1 CRFLKFLAKLKC 12

Scoring table: BLOSUM62

Searched: 449819 seqs, 56306365 residues

Database : Pending\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/paa/PCTUS9\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US084A\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084B\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	12	15	US-09-124-280A-20
2	50	73.5	10	15	US-09-124-280A-8
3	50	73.5	11	15	US-09-124-280A-37
4	45	66.2	10	15	US-09-124-280A-41
5	38	55.9	10	15	US-09-124-280A-10
6	38	55.9	635	16	US-09-248-796-16944
7	38	55.9	635	18	US-60-096-409-16944
8	37	54.4	752	1	PCT-US95-15138-21
9	37	54.4	752	7	US-08-420-235-21
10	37	54.4	752	7	US-08-420-235A-21
11	37	54.4	752	9	US-08-592-963B-21
12	37	54.4	752	21	US-08-793-624A-21
13	37	54.4	153	22	US-09-363-327-21
14	37	54.4	752	22	US-09-270-767-57424
15	36.5	53.7	337	22	US-09-270-767-57424
16	36	52.9	470	10	US-08-687-691A-4
17	36	52.9	552	16	US-09-248-796-25337
18	36	52.9	74	16	US-09-248-796-25684
19	36	52.9	552	18	US-60-096-409-25337
20	36	52.9	74	18	US-60-096-409-25684

21	36	52.9	454	19	PCT-US99-17130-298
22	36	52.9	314	22	US-09-270-767-45623
23	36	52.9	297	22	US-09-270-767-61133
24	35	51.5	1279	1	PCT-US98-06371-1122
25	35	51.5	155	1	PCT-US99-00338-18
26	35	51.5	1279	13	US-08-902-615A-388
27	35	51.5	7	15	US-09-124-280A-7
28	35	51.5	7	15	US-09-124-280A-38
29	35	51.5	155	16	US-09-225-585-18
30	35	51.5	74	16	US-09-248-796-26450
31	35	51.5	62	17	US-09-328-352-4453
32	35	51.5	225	17	US-09-328-352-5504
33	35	51.5	74	18	US-60-096-409-26450
34	35	51.5	67	22	US-09-270-767-35751
35	35	51.5	427	22	US-09-270-767-43924
36	35	51.5	132	22	US-09-270-767-44785
37	35	51.5	233	22	US-09-270-767-45076
38	35	51.5	67	22	US-09-270-767-50968
39	35	51.5	138	22	US-09-270-767-59330
40	34.5	50.7	344	7	US-08-446-919-4
41	34.5	50.7	365	14	US-09-077-173A-2
42	34	50.0	67	1	PCT-US98-11422A-240
43	34	50.0	66	1	PCT-US99-01621-104
44	34	50.0	376	1	PCT-US99-08605-2
45	34	50.0	680	11	US-08-761-136-1

ALIGNMENTS

RESULT 1  
US-09-124-280A-20  
; Sequence 20, Application US/09124280A  
; GENERAL INFORMATION:  
; APPLICANT: Porto, Massimo  
; TITLE OF INVENTION: VACCINES FOR PREVENTION OF GRAM-  
; TITLE OF INVENTION: NEGATIVE BACTERIAL INFECTIONS AND  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/124,280A  
; FILING DATE: July 29,1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-09-124-280A-20

Query Match 100.0%; Score 68; DB 15; Length 12;



ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/420,235B  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 45185-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 752 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-420-235B-21

Query Match 54.4%; Score 37; DB 2; Length 752;  
Best Local Similarity 41.7%; Pred. No. 89;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKFLKFLKC 12  
|:|:|:|:|:|  
Db 165 CRYLEFWEFEQC 176

RESULT 14  
PCT-US93-08528-188  
; Sequence 188, Application PC/TUS9308528  
; GENERAL INFORMATION:  
; APPLICANT: New York University  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/08528  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 188:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US93-08528-188

Query Match 54.4%; Score 37; DB 3; Length 32;  
Best Local Similarity 70.0%; Pred. No. 6.2;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FLKFLKFLKC 12  
|:|:|:|:|:|  
Db 15 FLLFLSLFNC 24

RESULT 15  
PCT-US95-10194-21  
; Sequence 21, Application PC/TUS9510194  
; GENERAL INFORMATION:  
; APPLICANT: The Trustees of Columbia University in the City of New York  
; APPLICANT: City  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/10194  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 45185-C-PCT/JPW/MS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 752 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-10194-21

Query Match 54.4%; Score 37; DB 3; Length 752;  
Best Local Similarity 41.7%; Pred. No. 89;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKFLKFLKC 12  
|:|:|:|:|:|  
Db 165 CRYLEFWEFEQC 176

Search completed: September 7, 1999, 22:38:39  
Job time: 7923 sec

us-09-124-280a-20.ra1

Wed Sep 8 09:42:59 1999

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-10

Query Match 55.9%; Score 38; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 1.7;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
DB 1 KFFKFFKFFK 10

RESULT 11  
US-08-456-112B-10  
Sequence 10, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porzio, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-10

Query Match 55.9%; Score 38; DB 2; Length 10;  
Best Local Similarity 70.0%; Pred. No. 1.7;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
DB 1 KFFKFFKFFK 10

RESULT 12  
US-08-118-270-188  
Sequence 188, Application US/08118270  
Patent No. 5508384  
GENERAL INFORMATION:  
APPLICANT: Murphy, Randall B.  
APPLICANT: Schuster, David I.  
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
TITLE OF INVENTION: RECEPTIONS, AND COMPOSITIONS AND METHODS THEREOF  
TITLE OF INVENTION: RECEPTIONS, AND COMPOSITIONS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118.270  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 188:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-188

Query Match 54.4%; Score 37; DB 1; Length 32;  
Best Local Similarity 70.0%; Pred. No. 6.2;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FLKFLKFLK 12  
DB 15 FLFLFLFLNC 24

RESULT 13  
US-08-420-235B-21  
Sequence 21, Application US/08420235B  
Patent No. 5801042  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-10

Query Match 55.9%; Score 38; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 1.7;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
DB 1 KFFKFFKFFK 10

RESULT 11  
US-08-456-112B-10  
Sequence 10, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porzio, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-10

Query Match 55.9%; Score 38; DB 2; Length 10;  
Best Local Similarity 70.0%; Pred. No. 1.7;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
DB 1 KFFKFFKFFK 10

RESULT 12  
US-08-118-270-188  
Sequence 188, Application US/08118270  
Patent No. 5508384  
GENERAL INFORMATION:  
APPLICANT: Murphy, Randall B.  
APPLICANT: Schuster, David I.  
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
TITLE OF INVENTION: RECEPTIONS, AND COMPOSITIONS AND METHODS THEREOF  
TITLE OF INVENTION: RECEPTIONS, AND COMPOSITIONS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118.270  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 188:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-188

Query Match 54.4%; Score 37; DB 1; Length 32;  
Best Local Similarity 70.0%; Pred. No. 6.2;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FLKFLKFLK 12  
DB 15 FLFLFLFLNC 24

RESULT 13  
US-08-420-235B-21  
Sequence 21, Application US/08420235B  
Patent No. 5801042  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

SEQUENCE CHARACTERISTICS:  
 LENGTH: 10 amino acids  
 TYPE: amino acid  
 TOPOLOGY: circular  
 US-08-456-112B-8

Query Match 73.5%; Score 50; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.029;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11  
 Db 1 KFLKFLKFLK 10

## RESULT 8

US-08-456-112B-37  
 ; Sequence 37, Application US/08456112B  
 ; Patent No. 5834430

GENERAL INFORMATION:  
 APPLICANT: Porro, Massimo  
 TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hedman, Gibson & Costigan  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 COMPUTER: LEADING EDGE 486  
 OPERATING SYSTEM: DOS  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/456,112B  
 FILING DATE: May 31, 1995  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:

FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Costigan, James V.  
 REGISTRATION NUMBER: 25,669  
 REFERENCE/DOCKET NUMBER: 576-004  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 302-8989  
 TELEFAX: (212) 302-8998  
 INFORMATION FOR SEQ ID NO: 37:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 11 amino acids  
 TYPE: amino acid  
 TOPOLOGY: circular  
 US-08-456-112B-37

Query Match 73.5%; Score 50; DB 2; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.032;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11  
 Db 2 KFLKFLKFLK 11

## RESULT 9

US-08-456-112B-41  
 ; Sequence 41, Application US/08456112B  
 ; Patent No. 5834430  
 GENERAL INFORMATION:  
 APPLICANT: Porro, Massimo

TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hedman, Gibson & Costigan  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 COMPUTER: LEADING EDGE 486  
 OPERATING SYSTEM: DOS  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/456,112B  
 FILING DATE: May 31, 1995  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:

FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Costigan, James V.  
 REGISTRATION NUMBER: 25,669  
 REFERENCE/DOCKET NUMBER: 576-004  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 302-8989  
 TELEFAX: (212) 302-8998  
 INFORMATION FOR SEQ ID NO: 41:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 10 amino acids  
 TYPE: amino acid  
 TOPOLOGY: circular  
 US-08-456-112B-41

Query Match 66.2%; Score 45; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFL 10  
 Db 2 KFLKFLKFL 10

## RESULT 10

US-08-097-830E-10  
 ; Sequence 10, Application US/08097830E  
 ; Patent No. 5652211

GENERAL INFORMATION:  
 APPLICANT: Porro, Massimo  
 TITLE OF INVENTION: Peptides For Neutralizing The  
 TOXICITY OF LIPID A  
 NUMBER OF SEQUENCES: 35  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: DOS  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/097,830E  
 FILING DATE:  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:

us-09-124-280a-20.ra1

Wed Sep 8 09:42:59 1999

QY 2 KFLKFLKFLK 11  
 Db 2 KFLKFLKFLK 11

RESULT 5

US-08-280-397-8  
 Sequence 8, Application US/08280397

Patent No. 5589459

GENERAL INFORMATION:

APPLICANT: Porro, Massimo  
 TITLE OF INVENTION: Synthetic Peptides for Detoxification  
 TITLE OF INVENTION: of Bacterial Endotoxins and for the  
 TITLE OF INVENTION: Prevention and Treatment of Septic  
 TITLE OF INVENTION: Shock  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/280,397

FILING DATE: 07/26/94

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/819,893

FILING DATE: 01/16/92

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-002A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 8:

LENGTH: 11 amino acids

TYPE: amino acid

TOPOLOGY: circular

US-08-280-397-8

Query Match 73.5%; Score 50; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.032; 0; Indels 0;  
 Matches 10; Conservative 0; Mismatches 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 Db 2 KFLKFLKFLK 11

RESULT 6

US-08-097-830E-8

Sequence 8, Application US/08097830E

Patent No. 5652211

GENERAL INFORMATION:

APPLICANT: Porro, Massimo  
 TITLE OF INVENTION: Peptides For Neutralizing The  
 TITLE OF INVENTION: Toxicity of Lipid A  
 NUMBER OF SEQUENCES: 35  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York

COUNTRY: USA  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: DOS  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/097,830E  
 FILING DATE:  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Costigan, James V.  
 REGISTRATION NUMBER: 25,669  
 REFERENCE/DOCKET NUMBER: 576-003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 302-8989  
 TELEFAX: (212) 302-8998  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 10 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-097-830E-8

Query Match 73.5%; Score 50; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.029; 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11

Db 1 KFLKFLKFLK 10

RESULT 7

US-08-456-112B-8

Sequence 8, Application US/08456112B

Patent No. 5834430

GENERAL INFORMATION:

APPLICANT: Porro, Massimo

TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson & Costigan

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: LEADING EDGE 486

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,112B

FILING DATE: May 31, 1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 8:

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; FILING DATE: May 31, 1995  
; APPLICATION NUMBER: US/08/456,112B  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-20

Query Match 100.0%; Score 68; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 7.9e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKFLKFLK 12  
Db 1 CKFLKFLKFLK 12

RESULT 3  
US-08-049-871-8  
; Sequence 8, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2 DOS  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-8

Query Match 73.5%; Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11  
Db 2 KFLKFLKFLK 11

RESULT 4  
US-07-819-893-8  
; Sequence 8, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
US-07-819-893-8

Query Match 73.5%; Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 17, Appl  
Sequence 9, Appl  
Sequence 17, Appl  
Sequence 18, Appl  
Sequence 42, Appl

40 31 45.6 87 1 US-07-847-743B-17  
41 31 45.6 351 1 US-08-196-350-1  
42 31 45.6 192 1 US-08-208-008C-9  
43 31 45.6 87 1 US-08-456-201-17  
44 31 45.6 10 1 US-08-097-830E-18  
45 31 45.6 41 1 US-08-168-091A-42

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:39 : Search time 80.79 seconds  
(without alignments)  
1.466 Million cell updates/sec

Title: US-09-124-280A-20  
Perfect score: 68  
Sequence: 1 CKFLKFLKLC 12

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	12	1	US-08-097-830E-20
2	68	100.0	12	2	US-08-456-112B-20
3	50	73.5	11	1	US-08-049-871-8
4	50	73.5	11	1	US-07-819-893-8
5	50	73.5	11	1	US-08-280-397-8
6	50	73.5	10	1	US-08-097-830E-8
7	50	73.5	10	2	US-08-456-112B-8
8	50	73.5	11	2	US-08-456-112B-37
9	45	66.2	10	2	US-08-456-112B-41
10	38	55.9	10	2	US-08-097-830E-10
11	38	55.9	10	2	US-08-456-112B-10
12	37	54.4	32	1	US-08-118-270-188
13	37	54.4	752	2	US-08-420-235B-21
14	37	54.4	752	3	PCT-US93-08528-188
15	37	54.4	752	3	PCT-US95-10194-21
16	35	51.5	7	1	US-07-819-893-9
17	35	51.5	7	1	US-08-280-397-9
18	35	51.5	7	1	US-08-097-830E-7
19	35	51.5	7	2	US-08-456-112B-7
20	35	51.5	7	2	US-08-456-112B-38
21	34.5	50.7	344	1	US-08-446-919A-4
22	34	50.0	680	2	US-08-211-430-2
23	33	48.5	207	1	US-07-656-566-2
24	33	48.5	231	1	US-07-656-566-3
25	33	48.5	299	1	US-08-118-270-77
26	33	48.5	12	2	US-08-218-026-25
27	33	48.5	12	2	US-08-218-026-40
28	33	48.5	12	2	US-08-653-632-25
29	33	48.5	12	2	US-08-653-632-40
30	33	48.5	9	2	US-08-456-112B-40
31	33	48.5	299	3	PCT-US93-08528-77
32	32	47.1	550	1	US-08-121-057-4
33	32	47.1	361	1	US-08-383-750-4
34	32	47.1	637	1	US-08-350-325A-10
35	32	47.1	550	2	US-08-509-187D-4
36	32	47.1	361	3	PCT-US93-09636-4
37	32	47.1	550	3	PCT-US93-09704A-4
38	32	47.1	2938	3	PCT-US94-00198-3
39	32	47.1	637	3	PCT-US94-03856-10

ALIGNMENTS

RESULT 1  
US-08-097-830E-20  
; Sequence 20, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-20

Query Match 100.0%; Score 68; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 7.9e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKFLKLC 12  
Db 1 CKFLKFLKLC 12

RESULT 2  
US-08-456-112B-20  
; Sequence 20, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45

---

**THIS PAGE BLANK (USPTO)**

Wed Sep 8 09:42:58 1999

fragments can be used in compositions for treating subjects suffering from a pathology related to a GPR abnormality e.g. a psychotic disorder such as schizophrenia.  
Sequence 32 AA;

Query Match 54.4%; Score 37; DB 1; Length 32;  
Best Local Similarity 70.0%; Pred. No. 7.9;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FLKFLKFLKC 12  
||| ||| |||  
Db 15 FLLEFLSFLNC 24

RESULT 13  
R93610  
ID R93610 standard; Protein; 752 AA.

AC R93610;  
DT 13-AUG-1996 (first entry)  
DE Kaposi's sarcoma associated herpesvirus ORF24 product.  
KW Kaposi's sarcoma; KSHV; lymphoma; AIDS; vaccine; diagnosis;  
KW therapy.  
OS Kaposi's sarcoma associated herpesvirus.  
PN WO9606159-A1.  
PD 29-FEB-1996.  
PF 11-AUG-1995; U10194.  
PR 18-AUG-1994; US-292365.  
PR 21-NOV-1994; US-343100.  
PR 11-APR-1995; US-420235.  
PA (UYCO ) UNIV COLUMBIA NEW YORK.  
PA (GRAM/) GRANT D E.  
PA (VIEL/) VIELE L.  
PI Chang Y, Moore PS, Grant DE, Viele L;  
DR WPI: 96-151362/15.  
DR N-PSDB: T16816.

PT Herpesvirus DNA associated with Kaposi's sarcoma - also associated  
PT vectors and proteins, used in detection and vaccination.  
PS Claim 17; Page 220-222; 305pp; English.  
CC Kaposi's sarcoma associated herpes virus (KSHV) clone KS5 (T16806),  
CC obtd. from a KS lesion genomic library, includes 15 complete ORFs and  
CC 2 incomplete ORFs (T16807-23) named according to their herpesvirus  
CC salmieri positional homologues. The ORF24 (T16816) product is  
CC given in R93610. KSHV proteins and peptides may be obtd. by  
CC incorporating encoding sequences into a vector and expression in  
CC host cells. They are useful in vaccines or for raising antibodies of  
CC diagnostic or therapeutic value.  
SQ Sequence 752 AA;

Query Match 54.4%; Score 37; DB 1; Length 752;  
Best Local Similarity 41.7%; Pred. No. 1.2e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKFLKFLKC 12  
|::|:|:|:|  
Db 165 CRYLEFWEFEQC 176

RESULT 14  
R97834  
ID R97834 standard; Protein; 752 AA.

AC R97834;  
DT 11-SEP-1996 (first entry)  
DE Kaposi's sarcoma associated herpesvirus ORF24 product.  
KW Kaposi's sarcoma; gamma-2 herpesvirus; KSHV; therapy; diagnosis;  
KW vaccine; diagnosis; AIDS.  
OS Kaposi's sarcoma associated herpesvirus.  
PN WO9615779-A1.  
PD 30-MAY-1996.  
PF 21-NOV-1995; U15138.  
PR 21-NOV-1994; US-343101.  
PR 11-APR-1995; US-420235.

(UYCO ) UNIV COLUMBIA NEW YORK.  
PA Chang Y, Moore PS;  
DR WPI: 96-268320/27.  
DR N-PSDB: T30686.  
PT Herpes virus associated with Kaposi's sarcoma - also definitive DNA  
PT sequences, useful for diagnosis of and to develop prods. for  
PT treatment of Kaposi's sarcoma  
PS Claim 17; Page 207-209; 277pp; English.  
CC Lambda clone KS5 (T30681) is a fragment of a newly identified human  
CC gamma-2 herpesvirus associated with Kaposi's sarcoma (KS). KS5 has  
CC 17 open reading frames (T30682-98), 15 of which are complete,  
CC including ORF24 (T30686). The protein products (R97830-46,  
CC respectively) of the 17 ORFs can be expressed in eukaryotic or  
CC bacterial host cells for use as vaccines, for KS diagnosis, or for  
CC raising antibodies.  
SQ Sequence 752 AA;

Query Match 54.4%; Score 37; DB 1; Length 752;  
Best Local Similarity 41.7%; Pred. No. 1.2e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKFLKFLKC 12  
|::|:|:|:|  
Db 165 CRYLEFWEFEQC 176

RESULT 15  
W02839  
ID W02839 standard; peptide; 32 AA.

AC W02839;  
DT 20-SEP-1996 (first entry)  
DE G-protein coupled receptor TM3 consensus polypeptide #94.  
KW G-protein coupled receptor; ligand binding assay; transmembrane domain;  
KW schizophrenia; dopamine; CAMP; adenosine; thrombin; adrenergic; opsin;  
KW muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin;  
KW odorant; cytomagalovirus; serotonergic.  
OS Synthetic.  
PN US5508384-A.  
PD 16-APR-1996.  
PF 10-SEP-1992; 943236.  
PR 10-SEP-1992; US-943236.  
PR 09-SEP-1993; US-118270.  
PA (UYNY ) UNIV NEW YORK STATE.  
PI Murphy RB, Schuster DI;  
DR WPI: 96-208785/21.  
DR New dopamine receptor peptide - useful as antipsychotic agent, e.g.  
PT for treating schizophrenia  
PT Disclosure; Column 235-236; 184pp; English.  
PS Polypeptides W02747-W02910 are based on the consensus transmembrane  
CC domain III sequence from G-protein coupled receptor (GPR) proteins  
CC and can be used in GPR ligand binding assays. The assays can be used  
CC to identify fragments, pref. transmembrane fragments, from GPR proteins  
CC (see W02657-W02635) which retain biological activity such as binding a  
CC GPR ligand or modulating a GPR ligand binding to a GPR (see W02747-W02999  
CC for examples of polypeptide fragments). The polypeptide fragments can be  
CC used in compsns. for treating subjects suffering from a pathology related  
CC to a GPR abnormality e.g. schizophrenia.  
SQ Sequence 32 AA;

Query Match 54.4%; Score 37; DB 1; Length 32;  
Best Local Similarity 70.0%; Pred. No. 7.9;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FLKFLKFLKC 12  
||| ||| |||  
Db 15 FLLEFLSFLNC 24

Search completed: September 7, 1999, 20:37:22  
Job time: 18471 sec



KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 45; Page 29; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 66.2%; Score 45; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 |||||  
 DB 2 KFLKFLKFL 10

RESULT 10  
 R71781  
 ID R71781 standard; peptide; 10 AA.  
 AC R71781;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI: 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 12; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides. 10 AA;

Query Match 55.9%; Score 38; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 2.1;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 |||||  
 DB 2 KFLKFLKFLK 11

DB 1 KKKKKKKKK 10  
 |||||

RESULT 11  
 W21598  
 ID W21598 standard; peptide; 10 AA.  
 AC W21598;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #10.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 14; Page 25; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 55.9%; Score 38; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 2.1;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 |||||  
 DB 1 KKKKKKKKK 10

RESULT 12  
 R50647  
 ID R50647 standard; peptide; 32 AA.  
 AC R50647;  
 DT 08-MAY-1996 (first entry)  
 DE G-protein coupled receptor TM3 consensus polypeptide #93.  
 KW G-protein coupled receptor; ligand binding assay; transmembrane domain;  
 KW psychotic disorder; schizophrenia; dopamine; CAMP; adenosine; thrombin;  
 KW muscarinic acetylcholine; adrenergic; endothelin; bombesin; endocrine;  
 KW rhodopsin; opsin; odorant; cytomegalovirus; serotonergic.  
 OS Synthetic.  
 PN WO9405695-A1.  
 PD 17-MAR-1994.  
 PF 09-SEP-1993; U08528.  
 PR 10-SEP-1992; US-943236.  
 PA (UNY ) UNIV NEW YORK STATE.  
 PI Murphy RB, Schuster DI;  
 DR WPI: 94-101120/12.  
 PT Polypeptides of G-coupled receptor proteins (GPRs) - useful for  
 PT binding GPR ligands or modulating GPR binding  
 PS Claim 9; Page 28; 160pp; English.  
 CC Polypeptides R48768-81 and R50569-R50718 are based on the consensus  
 CC transmembrane domain III sequence from G-protein coupled receptor (GPR)  
 CC proteins, and can be used in G-protein coupled receptor ligand binding  
 CC assays. The assay can be used to identify fragments pref. transmembrane  
 CC fragments, from GPR proteins (see R48686-R48758 for examples) which  
 CC retain biological activity such as binding a GPR ligand or modulating GPR  
 CC ligand binding to a GPR (see R48759-R48758, R50569-R50807 and  
 CC R89189-R89195 for examples of polypeptide fragments). The polypeptide

us-09-124-280a-20.rag

Wed Sep 8 09:42:58 1999

CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products. 11 AA;  
 CC Sequence 11 AA;

Query Match 73.5%; Score 50; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 DB 2 KFLKFLKFLK 11

RESULT 6  
 ID R71779 standard; peptide; 10 AA.  
 AC R71779;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M.  
 DR WPI; 95-075190/10.  
 DE New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 10; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides. 10 AA;  
 CC Sequence 10 AA;

Query Match 73.5%; Score 50; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.031;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 DB 1 KFLKFLKFLK 10

RESULT 7  
 ID W21625 standard; peptide; 11 AA.  
 AC W21625;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #37.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1995; E02313.  
 PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 41; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 CC Sequence 11 AA;

Query Match 73.5%; Score 50; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 DB 2 KFLKFLKFLK 11

RESULT 8  
 ID W21596 standard; peptide; 10 AA.  
 AC W21596;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #8.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1995; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 DE Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 12; Page 24; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 CC Sequence 10 AA;

Query Match 73.5%; Score 50; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.031;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 DB 1 KFLKFLKFLK 10

RESULT 9  
 ID W21629 standard; peptide; 10 AA.  
 AC W21629;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #41.

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 68; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CKFLKFLKFLKC 12  
 Db 1 CKFLKFLKFLKC 12

## RESULT 3

ID R33514 standard; peptide; 32 AA.  
 AC R33514;  
 DT 07-JUL-1993 (first entry)  
 DE Generic peptide for treating septic shock.  
 KW Cytokine shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-094304/11.  
 DR New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 2; Page 32; 39pp; English.  
 CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X  
 CC where n = 1-10, the peptide is useful for treating or preventing  
 CC septic shock, mixing with polymyxin B to reduce its toxicity;  
 CC removing endotoxins from blood, sera or other fluids (in vivo  
 CC or in vitro); controlling release of cytokines induced by  
 CC endotoxins; as diagnostic reagents to detect and quantify toxins  
 CC in blood or sera; preparing non-toxic antigenic complexes of  
 CC lipid A or lipopolysaccharide (LPS); and for treating pertussis,  
 CC bacterial meningitis and HIV-related infections. The usual dose is  
 CC 10-100 ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 32 AA;

Query Match 73.5%; Score 50; DB 1; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 0.084;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KFLKFLKFLK 11  
 Db 2 KFLKFLKFLK 11

## RESULT 4

ID R33532 standard; peptide; 11 AA.  
 AC R33532;  
 DT 07-JUL-1993 (first entry)

DE Peptide for treating septic shock.  
 KW Cytokine shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-094304/11.  
 DR New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 10; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His-Phe/Tyr/Trp-Leu/Ile/Val)n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 11 AA;

Query Match 73.5%; Score 50; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KFLKFLKFLK 11  
 Db 2 KFLKFLKFLK 11

## RESULT 5

ID R39294 standard; peptide; 11 AA.  
 AC R39294;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 DR New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 10; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic

Wed Sep 8 09:42:58 1999

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:21 ; Search time 147.16 Seconds  
(without alignments)  
1.931 Million cell updates/sec

Title: US-09-124-280A-20  
Perfect score: 68  
Sequence: 1 CKFLKFLKFLKC 12

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	68	100.0	12	1 R71791	Peptide neutralisi
2	68	100.0	12	1 W21608	Antibiotic potenti
3	50	73.5	32	1 R33514	Generic peptide fo
4	50	73.5	11	1 R33532	Peptide for treati
5	50	73.5	11	1 R33294	Endotoxin lipid A
6	50	73.5	11	1 R71779	Peptide neutralisi
7	50	73.5	11	1 W21625	Antibiotic potenti
8	50	73.5	11	1 W21596	Antibiotic potenti
9	45	66.2	10	1 W21629	Antibiotic potenti
10	38	55.9	10	1 W21598	Peptide neutralisi
11	38	55.9	10	1 R50847	Antibiotic coupled
12	37	54.4	32	1 R93610	G-protein coupled
13	37	54.4	752	1 R97834	Kaposi's sarcoma a
14	37	54.4	752	1 W02839	G-protein coupled
15	37	54.4	32	1 W02839	DAX-1 protein. New
16	36	52.9	470	1 W12678	Amphiphilic cation
17	36	52.9	19	1 W99418	Amphiphilic cation
18	36	52.9	19	1 W99419	Peptide for treati
19	35	51.5	7	1 R33533	Endotoxin lipid A
20	35	51.5	7	1 R39295	Peptide neutralisi
21	35	51.5	7	1 R71778	Antibiotic potenti
22	35	51.5	7	1 W21626	Antibiotic potenti
23	35	51.5	7	1 W21595	Antibiotic potenti
24	35	51.5	1279	1 W98761	H. pylori GHPO 109
25	34.5	50.7	365	1 W23605	Human P2Y4 recepto
26	34	50.0	680	1 R34445	Sequence encoded b
27	34	50.0	680	1 W49015	Human KAL protein.
28	34	50.0	67	1 W98540	Secreted protein e
29	34	50.0	67	1 W98019	Mouse calcium acti
30	34	50.0	425	1 W98019	Breast cancer asso
31	33	48.5	95	1 Y07002	Recombinant rat pr
32	33	48.5	225	1 P82078	Growth hormone-lik
33	33	48.5	232	1 R13583	Rat prolactin. Mas
34	33	48.5	226	1 R14599	Rat prolactin. Hyb
35	33	48.5	198	1 R22494	Human cytomagalovi
36	33	48.5	299	1 R48756	Human cytomagalovi
37	33	48.5	299	1 W02728	Antibiotic potenti
38	33	48.5	9	1 W21628	Follistatin relate
39	33	48.5	308	1 W17860	Prolactin antagoni
40	33	48.5	193	1 W23619	Prolactin antagoni
41	33	48.5	23	1 W23628	C. felis esterase,
42	33	48.5	550	1 W57865	C. felis esterase,
43	33	48.5	550	1 W57853	C. felis esterase,

## ALIGNMENTS

### RESULT 1

R71791  
ID R71791 standard; peptide; 12 AA.  
AC R71791;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..12  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M.  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids  
PS Claim 22; Page 21; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which include units of formula: (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new peptides.  
CC Sequence 12 AA:

Query Match 100.0%; Score 68; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. NO. 6.8e-05; Mismatches 0; Indels 0; Gaps 0;  
Matches 12; Conservative 0;

QY 1 CKFLKFLKFLKC 12  
| | | | | | | | | | | | | |  
Db 1 CKFLKFLKFLKC 12

### RESULT 2

W21608  
ID W21608 standard; peptide; 12 AA.  
AC W21608;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #20.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..12  
PN W09638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M; Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic required  
PS Claim 24; Page 26; 37pp; English.

C. felis esterase,  
Human 5' EST seque

**THIS PAGE BLANK (USPTO)**

Wed Sep 8 09:42:52 1999

Query Match 61.3%; Score 38; DB 12; Length 79;  
 Best Local Similarity 60.0%; Pred. No. 8.2;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
 DB 28 CKSSSFFFC 37

RESULT 14  
 O28864 PRELIMINARY; PRT; 144 AA.

AC O28864; 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-JAN-1998 (TEMBLrel. 07, Last annotation update)  
 DE HYPOTHETICAL 17.3 KD PROTEIN.  
 GN AFI408.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 CC Archaeoglobus.  
 RN [1]  
 RP SEQUENCE FROM N.A. 4304 / ATCC 49558;  
 RC STRAIN=VC-16 / DSM 98049343.  
 RX MEDLINE: 98049343.  
 RA KLECH H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,  
 RA RICHARDSON D.L., KERLAVENUSH J., GRAHAM D.E., KIRPES N.C.,  
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
 RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,  
 RA OVERBECK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,  
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,  
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESSE C.R.,  
 RA VENTER J.C.;  
 RA "The complete genome sequence of the hyperthermophilic,  
 RT sulphate-reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997).  
 DR EMBL; AE001007; AAB89647.1; -.  
 DR TIGR; AFI408; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 144 AA; 17349 MW; 68CE82B0 CRC32;

Query Match 59.7%; Score 37; DB 1; Length 144;  
 Best Local Similarity 87.5%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFC 9  
 DB 41 KKKKFFFC 48

RESULT 15  
 Q27472 PRELIMINARY; PRT; 583 AA.

AC Q27472;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
 DE SIMILAR TO CYTOCHROME P450. NCBI GI: 1255383 (EC 1.14.14.1).  
 GN C26F1.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 CC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE: 94150718;  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,

RA GARDNER A., GREEN P., HAWKINS J., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN K., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPOAT J., WOHLDMAN P.; III of C.  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA GEISEL C., STELLYES L., BRADSHAW H.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; U53148; AAB37074.1; -.  
 DR PFAM; PF00067; p450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.  
 FT BINDING 485 485 HEME (BY SIMILARITY).  
 SQ SEQUENCE 583 AA; 67366 MW; AEB7DD4C CRC32;

Query Match 59.7%; Score 37; DB 5; Length 583;  
 Best Local Similarity 66.7%; Pred. No. 59;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 9  
 DB 557 CEKKSFFVF 565

Search completed: September 7, 1999, 20:34:52  
 Job time: 19748 sec

RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
faicpaurum.";  
RL Science 282:1126-1132(1998).  
DR EMBL: AE001403; AAC71901.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 2206 AA; 264100 MW; AEE134B0 CRC32;

Query Match 61.3%; Score 38; DB 5; Length 2206;  
Best Local Similarity 87.5%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFEF 9  
| | | | | | | |  
Db 1285 KKKTFFEF 1292

RESULT 10  
O04488 PRELIMINARY; PRT; 382 AA.  
AC O04488;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
DE SEQUENCE OF BAC F21M12 FROM ARABIDOPSIS THALIANA CHROMOSOME 1,  
DE COMPLETE SEQUENCE.  
GN F21M12.4.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA VYSOTSKAYA V.S., OSBORNE B.I., TORIUMI M., YU G., QUI O., SHEN Y.K.,  
RA ARAUJO R., AU M., RUEHLER E., CONWAY A.B., CONWAY A.R., DEWAR K.,  
RA FENG J., KIM C., KURTZ D., LI Y., SHIN P., SUN H., DAVIS R.W.,  
RA ECKER J.R., FEDERSPIEL N.A., THEOLOGIS A.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC000132; AAB60722.1; -;  
SQ SEQUENCE 382 AA; 43555 MW; D732D7FB CRC32;

Query Match 61.3%; Score 38; DB 10; Length 382;  
Best Local Similarity 66.7%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KKKKFFEF 10  
| | | | | | | |  
Db 334 KKKLLFFC 342

RESULT 11  
Q43694 PRELIMINARY; PRT; 139 AA.  
AC Q43694;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE ACTIN DEPOLYMERIZING FACTOR.  
GN ZMABP2.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
OC Poaceae; Zea.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. A188; TISSUE=POLLEN;

RA LOPEZ I., ANTHONY R.G., MACIVER S.K., JIONG C.J., KHAN S., WEEDS A.G.,  
RA HUSSEY P.J.;  
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(0).  
DR EMBL: X97725; CAA66310.1; -;  
DR MENDEL; 11142; ZEMMA:1044;3.  
DR PFAM: PF00241; Cofilin\_ADF; 1.  
DR PROSITE: PS00325; ACTIN\_DEPOLYMERIZING; 1.  
SQ SEQUENCE 139 AA; 16083 MW; 91D491EF CRC32;

Query Match 61.3%; Score 38; DB 10; Length 139;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFEF 9  
| | | | | | | |  
Db 80 CQKSKIFFF 88

RESULT 12  
Q9ZSK2 PRELIMINARY; PRT; 146 AA.  
AC Q9ZSK2;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DE ACTIN DEPOLYMERIZING FACTOR 6.  
GN ADF6.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA DONG C.H., CHUA N.H.;  
RT "Arabidopsis Actin Depolymerizing Factor 6.";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF108284; AAD09112.1; -;  
SQ SEQUENCE 146 AA; 16708 MW; AC083C48 CRC32;

Query Match 61.3%; Score 38; DB 10; Length 146;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFEF 9  
| | | | | | | |  
Db 87 CQKSKIFFF 95

RESULT 13  
Q89373 PRELIMINARY; PRT; 79 AA.  
AC Q89373;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE HYPOTHETICAL PROTEIN.  
GN A38R.  
OS Paramexium bursaria chlorella virus 1 (PBCV-1).  
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 95133167.  
RA LU Z., LI Y., ZHANG Y., KUTISH G.F., ROCK D.L., VAN ETTEN J.L.;  
RT "Analysis of 45 kb of DNA located at the left end of the chlorella  
RT virus PBCV-1 genome.";  
RL Virology 206:339-352(1995).  
DR EMBL: U42580; AAC96406.1; -;  
SQ SEQUENCE 79 AA; 9448 MW; EGAB56BB CRC32;

Wed Sep 8 09:42:52 1999

```

082617 ID O82617 PRELIMINARY; PRT; 162 AA.
AC O82617;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE T9A4.2 PROTEIN.
GN T9A4.2.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eudicotyledons;
OC eudicotyledons; Magnoliophyta;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV. COLUMBIA;
RC WASHU;
RT "The A. thaliana Genome Sequencing Project.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=CV. COLUMBIA;
RC ZIDANIC M., MCQUERRY Y., SMITH A.;
RA "The sequence of A. thaliana T9A4.";
RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=CV. COLUMBIA;
RC WATERSTON R.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF096373; AAC62805.1;
DR EMBL: AF096373; AAC62805.1;
SQ SEQUENCE 162 AA; 17859 MW; 67910E87 CRC32;

Query Match 62.9%; Score 39; DB 10; Length 162;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CKKKKFFFC 10
Db 66 COERKGYFFC 75

RESULT 7
ID Q21325 PRELIMINARY; PRT; 935 AA.
AC Q21325; Q94264;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HYPOTHETICAL 107.8 KD PROTEIN.
GN K08D10.1 OR K08F11.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RC GEISEL C., BRADSHAW H.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE OF 47-921 FROM N.A.
RP STRAIN=BRISTOL N2;
RC MURRAY J., WOHLDMANN P.;
RA Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL: U55857; AAA98035.1;
DR EMBL: U55857; AAA98035.1;
DR EMBL: U70855; RAB09160.1;
SQ SEQUENCE 935 AA; 107799 MW; 6670B941 CRC32;

Query Match 61.3%; Score 38; DB 5; Length 935;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 2 KKKKFFFC 8
Db 39 KKKKFFFC 45

RESULT 8
ID O76688 PRELIMINARY; PRT; 273 AA.
AC O76688;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE M01G5.4 PROTEIN.
GN M01G5.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RC MEDLINE; 94150718;
RX WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., JIER M., JOHNSTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMANN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RC GEISEL C., KRAMER J., TWYMAN B.;
RA "The sequence of C. elegans cosmid M01G5.";
RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RC WATERSTON R.;
RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF078786; AAC26942.1;
DR EMBL: AF078786; AAC26942.1;
SQ SEQUENCE 273 AA; 32747 MW; C018D1F1 CRC32;

Query Match 61.3%; Score 38; DB 5; Length 273;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KKKKFFFC 9
Db 10 KKKKFFFC 17

RESULT 9
ID O96205 PRELIMINARY; PRT; 2206 AA.
AC O96205;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE HYPOTHETICAL 264.1 KD PROTEIN.
GN PFB0560W.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
RN [1]
RN SEQUENCE FROM N.A.
RP GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,

```



RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";   
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA BLANCHARD M., BRADSHAW H.;  
 RN Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U70857; AAB09170.1; -.  
 DR PFAM; PF00014; Kunitz\_BPTI; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ; 1.  
 KW Serine protease inhibitor.  
 SQ SEQUENCE 208 AA; 24008 MW; B3A1E540 CRC32;

Query Match 66.1%; Score 41; DB 5; Length 208;  
 Best Local Similarity 60.0%; Pred. No. 5.8;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CKKKKFFFC 10  
 | :||:| |  
 Db 173 CHOKYVFC 182

RESULT 3  
 O20167 PRELIMINARY; PRT; 62 AA.  
 ID O20167;  
 AC O20167;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE ORF62.  
 OS Chlorella vulgaris.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlorellales;  
 OC Chlorellaceae; Chlorella.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97303241.  
 RA WAKASUGI T., NAGAI T., KAPOOR M., SUGITA M., ITO M., ITO S.,  
 RA TSUDZUKI J., NAKASHIMA K., TSUDZUKI T., SUZUKI Y., HAMADA A., OHTA T.,  
 RA INAMURA A., YOSHINAGA K., SUGIURA M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the  
 RT green alga Chlorella vulgaris: the existence of genes possibly  
 RT involved in chloroplast division.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 DR EMBL; AB001684; BAA20683.1; -.  
 KW Chloroplast.  
 SQ SEQUENCE 62 AA; 7756 MW; 4DA77E1C CRC32;

Query Match 62.9%; Score 39; DB 8; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 KKKFFFF 9  
 | :||:| |

Db 16 KKKFFFF 22  
 RESULT 4  
 O20190 PRELIMINARY; PRT; 55 AA.  
 ID O20190;  
 AC O20190;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE ORF55C.  
 OS Chlorella vulgaris.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlorellales;  
 OC Chlorellaceae; Chlorella.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97303241.  
 RA WAKASUGI T., NAGAI T., KAPOOR M., SUGITA M., ITO M., ITO S.,  
 RA TSUDZUKI J., NAKASHIMA K., TSUDZUKI T., SUZUKI Y., HAMADA A., OHTA T.,  
 RA INAMURA A., YOSHINAGA K., SUGIURA M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the  
 RT green alga Chlorella vulgaris: the existence of genes possibly  
 RT involved in chloroplast division.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 DR EMBL; AB001684; BAA20720.1; -.  
 KW Chloroplast.  
 SQ SEQUENCE 55 AA; 6811 MW; 372F4DB1 CRC32;

Query Match 62.9%; Score 39; DB 8; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 KKKFFFF 9  
 | :||:| |  
 Db 13 KKKFFFF 19

RESULT 5  
 Q41700 PRELIMINARY; PRT; 1641 AA.  
 ID Q41700;  
 AC Q41700;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TReMBLrel. 09, Last annotation update)  
 DE PROTEIN BINDING TO ENO12B PROMOTER.  
 GN ENBP1.  
 OS Vicia sativa (Spring vetch) (Tare).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 OC Vicia.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA CHRISTIANSEN A., HANSEN A.C., VIJN I., PALLISGARD N., LARSEN K.,  
 RA YANG W.C., BISSELTING T., MARCKER K.A., JENSEN E.O.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X95995; CAA65242.1; -.  
 DR MENDEL; 17287; VICsa; 2675:mn17287.  
 SQ SEQUENCE 1641 AA; 181240 MW; 0A0C247D CRC32;

Query Match 62.9%; Score 39; DB 10; Length 1641;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CKKKKFFFC 10  
 | :||:| |  
 Db 832 CKKKKYCEC 841

RESULT 6

Wed Sep 8 09:42:52 1999

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:51; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-19  
Perfect score: 62  
Sequence: 1 CKKKKFFFC 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

- Database : SPTREMBL10.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	67.7	194	10	O81866 arabidopsis
2	41	66.1	208	5	Q94164 caenorhabdi
3	39	62.9	62	8	O20167 chlorella v
4	39	62.9	55	8	O20190 chlorella v
5	39	62.9	1641	10	O41700 vicia sativ
6	39	62.9	162	10	O82617 arabidopsis
7	38	61.3	935	5	O21325 caenorhabdi
8	38	61.3	273	5	O76688 caenorhabdi
9	38	61.3	2206	5	O96205 plasmodium
10	38	61.3	382	10	O04488 arabidopsis
11	38	61.3	139	10	Q43694 zea mays (m
12	38	61.3	146	10	O92SK2 arabidopsis
13	38	61.3	79	12	O89373 paramecium
14	37	59.7	144	1	O28864 archaeoglob
15	37	59.7	583	5	O27472 caenorhabdi
16	37	59.7	93	5	Q17196 bombyx mori
17	37	59.7	523	10	O92NU6 lycopersico
18	36	58.1	386	2	O56581 vibrio angu
19	36	58.1	282	2	O84180 chlamydia t
20	36	58.1	362	2	O927W1 chlamydia p
21	36	58.1	243	3	O94014 candida alb
22	36	58.1	330	5	O16721 caenorhabdi
23	36	58.1	372	5	O44849 caenorhabdi
24	36	58.1	7829	5	O18559 caenorhabdi
25	36	58.1	4550	5	O77336 plasmodium
26	36	58.1	2485	5	O96134 plasmodium
27	36	58.1	333	5	O97139 dictyosteli
28	36	58.1	1096	5	O97257 plasmodium
29	36	58.1	1629	10	O04698 pisum sativ

30	58.1	1701	10	O82022 medicago tr
31	35.5	770	5	Q18225 caenorhabdi
32	35.5	592	10	O04096 arabidopsis
33	35	250	2	Q50315 mycoplasma
34	35	553	4	Q13517 homo sapien
35	35	3097	5	O61143 plasmodium
36	35	465	5	O44732 caenorhabdi
37	35	138	5	Q18819 caenorhabdi
38	35	433	5	O17991 caenorhabdi
39	35	1817	5	O96253 plasmodium
40	35	3933	5	O97239 plasmodium
41	35	1827	5	O97275 plasmodium
42	35	783	12	O36393 alcelaphine
43	35	488	12	O98172 molluscum c
44	34	527	2	O84089 chlamydia t
45	34	165	5	O17367 caenorhabdi

## ALIGNMENTS

RESULT 1					
O81866					
ID	O81866	PRELIMINARY;	PRT;	194 AA.	
AC	O81866;				
DT	01-NOV-1998 (TREMELrel. 08, Created)				
DT	01-NOV-1998 (TREMELrel. 08, Last sequence update)				
DT	01-NOV-1998 (TREMELrel. 08, Last annotation update)				
DE	LECTIN LIKE PROTEIN.				
GN	T16H5.210.				
OS	Arabisopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;				
OC	core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;				
OC	Arabidopsis.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	DE HAAN M., MARSE A.C., GRIVELL L.A., BANCROFT I., MEWES H.W.,				
RA	MAYER K., SCHUELLER C., BEVAN M.;				
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	EU ARABIDOPSIS SEQUENCING PROJECT;				
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AL024486; CAA19702.1;				
KW	Lectin.				
SQ	SEQUENCE 194 AA; 22480 MW; B00DA8B7 CRC32;				

Query Match 67.7%; Score 42; DB 10; Length 194;  
Best Local Similarity 70.0%; Pred. No. 3.7;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKKKFFFC 10  
||| |  
DB 46 CKKVFVFEC 55

RESULT 2					
O94164					
ID	O94164	PRELIMINARY;	PRT;	208 AA.	
AC	O94164;				
DT	01-FEB-1997 (TREMELrel. 02, Created)				
DT	01-FEB-1997 (TREMELrel. 02, Last sequence update)				
DT	01-MAY-1999 (TREMELrel. 10, Last annotation update)				
DE	SIMILARITY TO BPT/KUNITZ INHIBITOR DOMAIN.				
GN	C10G8.2.				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;				
OC	Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				

**THIS PAGE BLANK (USPTO)**

Wed Sep 8 09:42:51 1999

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M35027; G335368; -  
DR PIR; B42506; B42506.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 96 AA; 11579 MW; C9378612 CRC32;

Query Match 54.8%; Score 34; DB 1; Length 96;  
Best Local Similarity 44.4%; Pred. No. 20;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 CKKKKFFFF 9  
Db 81 CKKRQYMFY 89

RESULT 15  
E322\_ADECG STANDARD; PRT; 195 AA.  
AC P22229;  
DT 01-AUG-1991 (REL. 19, CREATED)  
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE EARLY E3 22.2 KD GLYCOPROTEIN.  
OS CANINE ADENOVIRUS TYPE 1 (STRAIN GLAXO).  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; ADENOVIRIDAE; MASTADENOVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91272490.  
RA DRAGULEV B.P.; SIRA S.; ABOUHAIDAR M.G.; CAMPBELL J.B.;  
RT "sequence analysis of putative E3 and fiber genomic regions of two  
RT strains of canine adenovirus type 1";  
RL VIROLOGY 183:298-305(1991).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M60937; G210039; -  
DR PIR; C40318; ERADD2.  
KW EARLY PROTEIN; GLYCOPROTEIN.  
FT CARBOHYD 20 20 POTENTIAL.  
FT CARBOHYD 61 61 POTENTIAL.  
FT CARBOHYD 76 76 POTENTIAL.  
FT CARBOHYD 88 88 POTENTIAL.  
FT CARBOHYD 126 126 POTENTIAL.  
FT CARBOHYD 139 139 POTENTIAL.  
SQ SEQUENCE 195 AA; 22232 MW; 1F1CC9FB CRC32;

Query Match 53.2%; Score 33; DB 1; Length 195;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 FFFFC 10  
Db 5 FFFFC 9

Search completed: September 7, 1999, 23:50:22  
Job time: 1961 sec

```

FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
SQ SEQUENCE 240 AA; 26465 MW; B46535B1 CRC32;

Query Match 56.5%; Score 35; DB 1; Length 240;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKKKFFFC 10
DB 69 KKNLFFFC 77

RESULT 12
MTHL_HAEN
ID MTHL_HAEN STANDARD; PRT; 359 AA.
AC P20590;
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DE MODIFICATION METHYLASE HINFI (SC 2.1.1.72) (ADENINE-SPECIFIC
DE METHYLTRANSFERASE HINFI) (M.HINFI).
GN HINFI.
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC HAEMOPHILUS.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=RF;
RC MEDLINE; 89108022.
RX CHANDRASEGARAN S., LUNNEN K.D., SMITH H.O., WILSON G.G.;
RT "Cloning and sequencing the HinfI restriction and modification
RT genes".
RL GENE 70:387-392(1988).
CC -----
CC -!- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC GATTC, CAUSES SPECIFIC METHYLATION ON A-2 ON BOTH STRANDS, AND
CC PROTECTS THE DNA FROM CLEAVAGE BY THE HINFI ENDONUCLEASE.
CC
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA ADENINE -
CC S-ADENOSYL-L-HOMOCYSTEINE + DNA 6-METHYLMINOPURINE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22862; G148945; -
CC PIR; JT0391; JT0391.
CC REBASE; RB00028; HinfI.
CC PROSITE; PS00092; N6_MTASE; 1.
KW TRANSFERASE; METHYLTRANSFERASE; RESTRICTION SYSTEM.
SQ SEQUENCE 359 AA; 41805 MW; B4EDCB0B CRC32;

Query Match 54.8%; Score 34; DB 1; Length 359;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKKKFFF 8
DB 148 CKKNKTF 155

RESULT 13
YHN4_YEAST
ID YHN4_YEAST STANDARD; PRT; 714 AA.
AC P38795;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)

```

```

DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 80.7 KD PROTEIN IN ERG7-NMD2 INTERGENIC REGION.
GN YHR074W.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C / AB972;
RC MEDLINE; 94378003.
RX JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
RA DU 2., FAVELLO A., FULTON L., GATTUNG S., GEISEL C., KIRSTEN J.,
RA KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
RA LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
RA NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASKIS E., VAUGHAN K.,
RA VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
RA VAUDIN M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL SCIENCE 265:2077-2082(1994).
CC -!- SIMILARITY: LOW, TO B.SUBTILIS NH(3)-DEPENDENT NAD(+) SYNTHETASE
CC (OUTB).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U10556; G500832; -
CC PIR; S46811; S46811.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 714 AA; 80685 MW; 11F0F7B6 CRC32;

Query Match 54.8%; Score 34; DB 1; Length 714;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKFFFF 9
DB 641 KVKRFFFF 648

RESULT 14
YVFC_VACCC
ID YVFC_VACCC STANDARD; PRT; 96 AA.
AC P20561;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 11.6 KD PROTEIN.
GN F ORF C.
OS VACCINIA VIRUS (STRAIN COPENHAGEN).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUS.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 91021027.
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RT "The complete DNA sequence of vaccinia virus."
RL VIROLOGY 179:247-266(1990).
[2]
RN COMPLETE GENOME.
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RL VIROLOGY 179:517-563(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```
Query Match      58.18; Score 36; DB 1; Length 428;
Best Local Similarity 77.88; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CKKKKFFFF 9
DB 80 CKKFFGFF 88

RESULT 10
CRL_ECOLI
ID CRL_ECOLI STANDARD; PRT; 132 AA.
AC P24251; P77629;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CURLIN GENES TRANSCRIPTIONAL ACTIVATORY PROTEIN.
GN CRL
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE: 83189086.
RA OVERBEKE N., BERGMAN H., VAN MANSFELD F., LUGTENBERG B.;
RT "complete nucleotide sequence of phoE, the structural gene for the
RT phosphate limitation inducible outer membrane pore protein of
RT Escherichia coli K12."
RT J. MOL. BIOL. 163:513-532(1983).
RN [2]
RP SEQUENCE FROM N.A., SEQUENCE OF 1-5, AND CHARACTERIZATION.
RC STRAIN-A012;
RX MEDLINE: 89201357.
RA OLSEN A., JONSSON A., NORMARK S.;
RT "Fibronectin binding mediated by a novel class of surface organelles
RT on Escherichia coli."
RL NATURE 338:652-655(1989).
RN [3]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN-A012;
RX MEDLINE: 93023873.
RA ARNOVIST A., OLSEN A., PFEIFER J., RUSSELL D.G., NORMARK S.;
RT "The Crl protein activates cryptic genes for curli formation and
RT fibronectin binding in Escherichia coli HB101."
RL MOL. MICROBIOL. 6:2443-2452(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE: 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL SCIENCE 277:1453-1474(1997).
RN [5]
RP SEQUENCE FROM N.A.
RA SCHRAMM S., DUNCAN M., ALLEN E., ARAUJO R., APARICIO A., CHUNG E.,
RA DAVIS K., FEDERSPIEL N., HYMAN R., KALMAN S., KOMP C., KURDI O.,
RA LASHKARI D., LEW H., LIN D., NAMATH A., OEFNER P., ROBERTS D.,
RA DAVIS R.W.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RA TAKEMOTO K., MORI H., MURAYAMA N., KATAOKA K., YANO M., ITOH T.,
RA YAMAMOTO Y., INOKUCHI H., MIKI T., HATADA E., FUKUDA R.,
RA ICHIHARA S., MIZUNO T., MAKINO K., NAKATA A., YURA T., SAMPEI G.,
RA MIZOBUCHI K.;
RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: BEHAVES AS A TRANSCRIPTIONAL MODULATOR ACTIVATING THE
CC GENE FOR CURLIN AND PROBABLY OTHER GENES INVOLVED IN CURLI
```

```
FORMATION.
-!- SUBCELLULAR LOCATION: CYTOPLASMIC.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL: V00316; G42392; --
EMBL: X67207; G41154; --
EMBL: AE000132; G1786435; --
EMBL: U70214; G1552808; --
EMBL: D83536; D1012656; --
PIR: A05115; A05115
ECOGENE; EG11092; CRL.
KW TRANSCRIPTION REGULATION; ACTIVATOR.
FT INIT_MET 0
FT CONFLICT 7 7 P -> L (IN REF. 1).
FT CONFLICT 12 12 I -> N (IN REF. 1).
FT CONFLICT 28 28 E -> K (IN REF. 1, 4 AND 5).
SQ SEQUENCE 132 AA; 15525 MW; 22612B1F CRC32;

Query Match      56.5%; Score 35; DB 1; Length 132;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 CKKKKFFFFC 10
DB 27 CEDNRFEEFC 36

RESULT 11
YCXN_PORPU
ID YCXN_PORPU STANDARD; PRT; 240 AA.
AC P51364;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 26.5 KD PROTEIN IN YCF44-CPG INTERGENIC REGION (ORF240).
OS PORPHYRA PURPUREA.
OC CHLOROPLAST.
OC EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; BANGIALES; BANGIACEAE; PORPHYRA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AVONPORT;
RA REITH M.E., MUNHOLLAND J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome."
RL PLANT MOL. BIOL. REP. 13:333-335(1995).
CC -!- FUNCTION: COULD BE INVOLVED IN CYTOCHROME C SYNTHESIS
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CHLOROPLAST
CC (PROBABLY).
CC -!- SIMILARITY: TO B. SUBTILIS CCDA AND H. INFLUENZAE H11454.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL: U38804; G1276830; --
HYPOTHETICAL PROTEIN; CYTOCHROME C-TYPE BIOGENESIS; TRANSMEMBRANE;
KW CHLOROPLAST.
FT TRANSMEM 32 52 POTENTIAL.
FT TRANSMEM 74 94 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
```

CC CEREBELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.  
CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L07868; G337360; -  
DR MIN; 600543; -  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PFAM; PF00069; kinase; 1.  
DR PFAM; PF00757; Furin-like; 1.  
DR PFAM; PF01030; Recept\_L\_Domain; 2.  
DR TRANSMEMBRANE; GLYCOPROTEIN; MULTIGENE FAMILY; RECEPTOR; SIGNAL;  
KW TRANSFERASE; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;  
KW ALTERNATIVE SPLICING.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 1308 ERBB-4 RECEPTOR PROTEIN-TYROSINE KINASE.  
FT DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 652 675 POTENTIAL.  
FT DOMAIN 676 1308 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 186 334 CYSTEINE-RICH.  
FT DOMAIN 496 633 CYSTEINE-RICH.  
FT DOMAIN 713 988 PROTEIN KINASE.  
FT NP\_BIND 724 732 ATP (BY SIMILARITY).  
FT BINDING 751 751 ATP (BY SIMILARITY).  
FT ACT\_SITE 843 843 BY SIMILARITY.  
FT MOD\_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT VARSPLIC 626 648 NGPTSHDCIVYPTMGHSTLPQHA -> IGSSIEDICGLMD  
(IN ISOFORM JM-B).  
FT CARBOHYD 138 138 POTENTIAL.  
FT CARBOHYD 174 174 POTENTIAL.  
FT CARBOHYD 181 181 POTENTIAL.  
FT CARBOHYD 253 253 POTENTIAL.  
FT CARBOHYD 358 358 POTENTIAL.  
FT CARBOHYD 410 410 POTENTIAL.  
FT CARBOHYD 473 473 POTENTIAL.  
FT CARBOHYD 495 495 POTENTIAL.  
FT CARBOHYD 548 548 POTENTIAL.  
FT CARBOHYD 576 576 POTENTIAL.  
FT CARBOHYD 620 620 POTENTIAL.  
SQ SEQUENCE 1308 AA; 146807 MW; 64A58682 CRC32;  
  
Query Match 81.1%; Score 30; DB 1; Length 1308;  
Best Local Similarity 71.4%; Pred. No. 1.8e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2 ELKRIKI 8  
Db 717 ELKRVKV 723  
-----  
RESULT 7  
ID MFAL\_CHICK STANDARD; PRT; 442 AA.  
AC P55080;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MICROFIBRILLAR-ASSOCIATED PROTEIN 1 (ASSOCIATED MICROFIBRIL PROTEIN)  
DE (AMF).  
GN MFAP1.  
OS GALLUS GALLUS (CHICKEN).  
-----

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=WHITE LEGHORN; TISSUE=EMBRYO;  
RX MEDLINE; 92250568.  
RT HERRIGAN S.K., RICH C.B., STREETEN B.W., LI Z.-Y., FOSTER J.A.;  
RT "Characterization of an associated microfibrillar protein through  
RT recombinant DNA techniques".  
RL J. BIOL. CHEM. 267:10087-10095(1992).  
CC -!- FUNCTION: COMPONENT OF THE ELASTIN-ASSOCIATED MICROFIBRILS.  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.  
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M76679; -; NOT\_ANNOTATED\_CDS.  
DR EXTRACELLULAR MATRIX.  
FT DOMAIN 137 142 POLY-GLU.  
SQ SEQUENCE 442 AA; 51964 MW; D2C7AA58 CRC32;  
  
Query Match 81.1%; Score 30; DB 1; Length 442;  
Best Local Similarity 85.7%; Pred. No. 62;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 KELKRIK 7  
Db 285 RELKRIK 291  
-----  
RESULT 8  
ID MFAL\_HUMAN STANDARD; PRT; 439 AA.  
AC P55081;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MICROFIBRILLAR-ASSOCIATED PROTEIN 1.  
GN MFAP1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SKIN;  
RX MEDLINE; 95137591.  
RX YEH H., CHOW M., ABRAMS W.R., FAN J., FOSTER J., MITCHELL H.,  
RA MUENKE M., ROSENBLUM J.;  
RT "Structure of the human gene encoding the associated microfibrillar  
RT protein (MFAP1) and localization to chromosome 15q15-q21".  
RL GENOMICS 23:443-449(1994).  
CC -!- FUNCTION: COMPONENT OF THE ELASTIN-ASSOCIATED MICROFIBRILS.  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U04209; G434656; -  
DR MIN; 600215; -  
KW EXTRACELLULAR MATRIX.  
FT DOMAIN 134 139 POLY-GLU.

cc or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

FT	DOMAIN	I	246		PRECORRIN-8X METHYLMUTASE.
FT	DOMAIN	247	446		
SO	SEQUENCE	446 AA;	48678 MW;	A379342E CRC32;	

Query Match	81.1%	Score 30;	DB 1;	Length 440;
Best Local Similarity	75.0%	Pred. No. 62;		
Matches 6;	Conservative	1;	Mismatches	1;
			Indels	0;
			Gaps	0

RESULT 6  
ERB4\_HUMAN  
ID ERB4 HUMAN  
STANDARD;  
PRT; 1308 AA.

AC	Q15303;
DT	15-DEC-1998 (REL. 37, CREATED)
DT	15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DT	15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE	ERBB-4 RECEPTOR PROTEIN-TYROSINE KINASE PRECURSOR (EC 2.7.1.112)
DE	(P180 ERB4).
OS	ERBB4 OR HER4

OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]

RP SEQUENCE FROM N.A. (ISOFORM JM-A).  
RC TISSUE-BREAST CARCINOMA;  
RX MEDLINE: 93189574.  
RA PLOWMAN G.D. CULOUSCOU J.-M., WHITNEY G.S., GREEN J.M., CARLTON G.W.,  
RA FLOY L., NEUBAUER M.G., SHOYAB M.:  
RA "Ligand-specific activation of HER4/p180erbB4, a fourth member of the  
RT epidermal growth factor receptor family".  
RT epidermal growth factor receptor family".  
RL PROC. NATL. ACAD. SCI. U.S.A. 90:1746-1750(1993).  
RN [2].  
RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).  
RC TISSUE-FETAL BRAIN;  
RX MEDLINE: 97476287.  
RA ELENIUS K., CORFAS G., PAUL S., CHOI C.J., RIO C., PLOWMAN G.D.,  
RA KLAGEBURN M.:  
RA "A novel juxtamembrane domain isoform of HER4/erbB4. Isoform-specific  
RT tissue distribution and differential processing in response to  
RT phorbol ester".  
RT J. BIOL. CHEM. 272:26761-26768(1997).

**CC**      **CC**      **CC**

**U:**    **BLO:**    **CRA:**    **CC**

**-!** - FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULIN, NRG-2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND DIFFERENTIATION.

NTAK. INTERACTION WITH THESE AGENTS IS NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN. -!- CATALYTIC ACTIVITY:  $ATP + A \text{ PROTEIN TYROSINE} = ADP +$

CC PROTEIN TYROSINE PHOSPHATE.  
CC -|- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
CC RESIDUES.

CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ENDBEADS.  
CC RECEPTORS (POTENTIAL).  
CC SUGGESTED REPLICATION: TYPE I MEMBRANE PROTEIN.

CC - SUBCELLULAR LOCATION: EXPRESSED AT HIGHEST LEVELS IN THYMUS,  
CC - TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN THYMUS,  
CC - KIDNEY, IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREBELLUM,  
CC - SUBCELLULAR LOCATION: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,  
CC - TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,  
CC - SUBCELLULAR LOCATION: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,

CC PITUITARY, SPLEEN, TESTIS AND PANCREAS.  
CC LUNG, SALIVARY GLAND, AND PANCREAS.  
CC -:- ALTERNATIVE PRODUCTS: TWO ISOFORMS: JM-A (SHOWN HERE) AND JM-B;  
CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER  
CC FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER: JM-A IS PROCESSED  
CC BUT NOT JM-B. SO, THEY RESPECTIVELY REPRESENT CLEAVABLE AND  
CC NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN



```

CC EMBL; J04739; G179529; ALT_INIT.
DR PIR; A29464; A29464.
DR PIR; A30909; A30909.
DR PIR; A33850; A33850.
DR PDB; 1BP1; 04-SEP-97.
DR MIN; 109195; -.
DR PROSITE; PS00400; LBP_BPI_CETP; 1.
DR PFAM; PF01273; LBP_BPI_CETP; 1.
KW ANTIMIOTIC; SIGNAL; TRANSMEMBRANE; GLYCOPROTEIN; 3D-STRUCTURE.
FT SIGNAL 1 27
FT CHAIN 28 483 BACTERICIDAL PERMEABILITY-INCREASING
FT SITE 236 241 CLEAVAGE SITES FOR ELASTASE (POTENTIAL).
FT TRNMEM 365 385 POTENTIAL.
FT SEQUENCE 483 AA; 53396 MW; 0AAEF65 CRC32;
SQ SEQUENCE 483 AA; 53396 MW; 0AAEF65 CRC32;

Query Match 100.0%; Score 37; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8
Db 54 KELKRIKI 61

RESULT 2
TRUB_AQUAE STANDARD; PRT; 287 AA.
AC 06922;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE TRNA PSEUDOURIDINE SYNTHASE B (EC 4.2.1.70) (TRNA PSEUDOURIDINE 55
DE SYNTHASE) (PSI55 SYNTHASE) (PSEUDOURIDYLATE SYNTHASE) (URACIL
DE HYDROLYASE).
GN TRUB.
OS AQUIFEX AEROLICUS.
OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5.
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RA "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL NATURE 392:353-358(1998).
CC -!- FUNCTION: FORMATION OF PSEUDOURIDINE AT POSITION 55 IN THE
CC ANTICODON STEM AND LOOP OF TRANSFER RNAs (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: URACIL + D-RIBOSE 5-PHOSPHATE -> PSEUDOURIDINE
CC 5'-PHOSPHATE + H(2)O.
CC -!- SIMILARITY: BELONGS TO THE TRUB FAMILY OF PSEUDOURIDINE SYNTHASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000703; G2983293; -.
DR EMBL; AE000703; G2983293; -.
KW LYASE; TRNA PROCESSING.
SQ SEQUENCE 287 AA; 32259 MW; 633369D0 CRC32;

```

```

Query Match 86.5%; Score 32; DB 1; Length 287;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KELKRIKI 8
Db 190 KELRLKI 197

RESULT 3
CGRI_CANAL STANDARD; PRT; 340 AA.
ID CGRI_CANAL
AC P56553;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CELL GROWTH PROTEIN CGRI.
GN CGRI.
OS CANDIDA ALBICANS (YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC CANDIDACEAE; CANDIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 9061;
RA CHO T., SUDOH M., TANAKA T., NAKASHIMA Y.;
RA SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: INVOLVED IN THE CELL GROWTH REGULATION.
CC -!- SIMILARITY: BELONGS TO THE SMP-30 / CGRI FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB013716; D1029116; -.
DR EMBL; AB013716; D1029116; -.
SQ SEQUENCE 340 AA; 38391 MW; 5B19EC42 CRC32;

Query Match 83.8%; Score 31; DB 1; Length 340;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8
Db 264 KEIERIKI 271

RESULT 4
CAS2_RABIT STANDARD; PRT; 180 AA.
ID CAS2_RABIT
AC P50418;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE ALPHA-S2A CASEIN PRECURSOR.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=MAMMARY GLAND;
RX MEDLINE; 94107245.
RA DAWSON S.P., WILDE C.J., TIGHE P.J., MAYER R.J.;
RA "Characterization of two novel casein transcripts in rabbit mammary
RT gland";
RL BIOCHEM. J. 296:777-784(1993).
CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CC CALCIUM PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:05 ; Search time 71.09 Seconds  
(without alignments)  
3.181 Million cell updates/sec

Title: US-09-124-280A-25  
Perfect score: 37  
Sequence: 1 KELRIKI 8

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	483	1	BPI_HUMAN
2	32	86.5	287	1	TRUE_AQUAE
3	31	83.8	340	1	CGR1_CANAL
4	30	81.1	180	1	CAS2_RABIT
5	30	81.1	446	1	COBJ_ARCFU
6	30	81.1	1308	1	ERB4_HUMAN
7	30	81.1	442	1	MFAL_CHICK
8	30	81.1	439	1	MFAL_HUMAN
9	30	81.1	2165	1	RRPL_HRSVA
10	30	81.1	1267	1	YML3_YEAST
11	30	81.1	1131	1	YPO4_CAEEL
12	29	78.4	4563	1	APB_HUMAN
13	29	78.4	715	1	DCLY_SALTY
14	29	78.4	1138	1	DPS2_YEAST
15	29	78.4	1829	1	MYSD_CHICK
16	29	78.4	1480	1	PAN1_YEAST
17	29	78.4	449	1	PS59_LYCES
18	29	78.4	211	1	Y290_METJA
19	29	78.4	486	1	YB10_METJA
20	29	78.4	186	1	YCF4_PORPU
21	29	78.4	427	1	YKR2_YEAST
22	28	75.7	372	1	3BH1_HUMAN
23	28	75.7	372	1	3BHS_MACMU
24	28	75.7	184	1	A412_PLAFA
25	28	75.7	909	1	CSKP_RAT
26	28	75.7	740	1	FAS_PNECA
27	28	75.7	358	1	GLPO_ECOLI
28	28	75.7	364	1	GLPO_HAEIN
29	28	75.7	356	1	GLPQ_TREPA
30	28	75.7	655	1	ILFI_HUMAN
31	28	75.7	1079	1	MYIA_MOUSE
32	28	75.7	458	1	NTPB_ENTHR
33	28	75.7	1089	1	PGDS_MOUSE
34	28	75.7	377	1	PGT1_HUMAN
35	28	75.7	377	1	PGT1_RAT
36	28	75.7	1088	1	PRL1_PETCR
37	28	75.7	254	1	P222_YEAST
38	28	75.7	1297	1	PURL_HAEIN
39	28	75.7	423	1	PYRC_METJA
40	28	75.7	101	1	RS14_ACYKS
41	28	75.7	479	1	S61A_SCHPO
42	28	75.7	471	1	S61A_YARLI
43	28	75.7	1189	1	SCII_CHICK

44	28	75.7	228	1	SODM_CAPAN	O49066 capsicum an
45	28	75.7	53	1	VGK_BPPHK	Q38038 bacterioph

ALIGNMENTS

RESULT 1

ID	BPI_HUMAN	STANDARD;	PRT;	483 AA.
AC	P17213;			
DT	01-AUG-1990 (REL. 15, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
DE	BACTERICIDAL PERMEABILITY-INCREASING PROTEIN PRECURSOR (BPI) (CAP 57).			
GN	BPI.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 28-64.			
RX	MEDLINE: 89255455.			
RA	GRAY P.W., FLAGGS G., LEONG S.R., GUMINA R.J., WEISS J., OOI C.E.,			
RA	ELSBACH P.;			
RT	"Cloning of the cDNA of a human neutrophil bactericidal protein.			
RT	Structural and functional correlations.;"			
RL	J. BIOL. CHEM. 264:9505-9509(1989).			
RN	[2]			
RP	SEQUENCE OF 28-42.			
RX	MEDLINE: 88033057.			
RA	OOI C.E., WEISS J., ELSBACH P., FRANGIONE B., MANNION B.;			
RT	"A 25-kDa NH2-terminal fragment carries all the antibacterial			
RT	activities of the human neutrophil 60-kDa			
RT	bactericidal/permeability-increasing protein.;"			
RL	J. BIOL. CHEM. 262:14891-14894(1987).			
RN	[3]			
RP	SEQUENCE OF 28-47.			
RX	MEDLINE: 89315847.			
RA	GABAY J.E., SCOTT R.W., CAMPANELLI D., GRIFFITH J., WILDE C.,			
RA	MARRA M.N., SEEGER M., NATHAN C.F.;			
RT	"Antibiotic proteins of human polymorphonuclear leukocytes.;"			
RL	PROC. NATL. ACAD. SCI. U.S.A. 86:5610-5614(1989).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).			
RX	MEDLINE: 97334442.			
RA	BEAMER L.J., CARROLL S.F., EISENBERG D.;			
RT	"Crystal structure of human BPI and two bound phospholipids at 2.4-A			
RT	resolution.;"			
RL	SCIENCE 276:1861-1864(1997).			
CC	-1- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES			
CC	OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A			
CC	STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE			
CC	NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE			
CC	GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR			
CC	LEUKOCYTES (PMN) GRANULES.			
CC	-1- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES.			
CC	-1- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF			
CC	THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE			
CC	MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE			
CC	RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-			
CC	AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED			
CC	RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE			
CC	INGESTED.			
CC	-1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			

**THIS PAGE BLANK (USPTO)**

Wed Sep 8 09:43:28 1999

A:Map position: 13R  
C:Superfamily: unassigned DEAD/H box helicases; DEAD/H box helicase homology  
C:Keywords: ATP; P-loop; transmembrane protein  
F:414-815/Domain: DEAD/H box helicase homology <DEAD>  
F:414-421/Region: nucleotide-binding motif A (P-loop)  
F:512-517/Region: nucleotide-binding motif B  
F:516-519/Region: DEAH motif  
F:922-938/Domain: transmembrane #status predicted <TMM>

Query Match 81.1%; Score 30; DB 2; Length 1267;  
Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KELKRIK 7  
Db 21 KELKRIR 27

Search completed: September 7, 1999, 23:22:40  
Job time: 947 sec

A:Reference number: A40317; MUID:91272488

A:Accession: A40317

A:Molecule type: mRNA

A:Residues: 1-2165 <STE>

A:Cross-references: GB:M75730; NID:g333955; PID:g333956

R:Collins, P.L.; Olmsted, R.A.; Spriggs, M.K.; Johnson, P.R.; Buckler-White, A.J.

Proc. Natl. Acad. Sci. U.S.A. 84, 5134-5138, 1987

A:Title: Gene overlap and site-specific attenuation of transcription of the viral polyome

A:Reference number: A28319; MUID:87260943

A:Accession: A28319

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-81 <COL>

A:Cross-references: GB:M17245; NID:g333953; PID:g333954

R:Johnson, P.R.; Collins, P.L.

J. Gen. Virol. 69, 2901-2906, 1988

A:Title: The A and B subgroups of human respiratory syncytial virus: comparison of inter

A:Reference number: PS0048; MUID:89036169

A:Accession: PS0048

A:Molecule type: mRNA

A:Residues: 1-18 <JOH>

A:Cross-references: GB:D00397; NID:g222551; PID:d1000750; PID:g2160375

A:Experimental source: Strain 18537

A:Note: this strain belongs to subgroup B

C:Genetics:

A:Gene: L

C:Superfamily: parainfluenza virus RNA-directed RNA polymerase

C:Keywords: ATP; nucleotidyltransferase

Query Match

Best Local Similarity 81.1%; Score 30; DB 1; Length 2165;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELKRIKI 8

|||||:

Db 256 KELKRIYI 263

RESULT 12

A47253

epidermal growth factor receptor, HER4 - human

C:Species: Homo sapiens (man)

C:Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998

C:Accession: A47253

R:Plowman, G.D.; Culouscou, J.M.; Whitney, G.S.; Green, J.M.; Carlton, G.W.; Foy, L.; Ne

Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993

A:Title: Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal

A:Reference number: A47253; MUID:93189574

A:Accession: A47253

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-1308 <ELO>

A:Cross-references: GB:L07868; NID:g337359; PID:g337360

A:Note: sequence extracted from NCBI backbone (NCBIP:126842)

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP; growth factor receptor

F:716-981/Domain: protein kinase homology <KIN>

F:724-732/Region: protein kinase ATP-binding motif

Query Match

Best Local Similarity 81.1%; Score 30; DB 2; Length 1308;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 ELKRIKI 8

|||||:

Db 717 ELKRVKV 723

RESULT 13

S39775

alpha-s2-casein form a precursor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 07-Oct-1994 #sequence\_revision 01-Dec-1995 #text\_change 21-Aug-1998

C:Accession: S39775

R:Dawson, S.P.; Wilde, C.J.; Tighe, P.J.; Mayer, R.J.

Biochem. J. 296, 777-784, 1993

A:Title: Characterization of two novel casein transcripts in rabbit mammary gland.

A:Reference number: S39775

A:Accession: S39775

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-180 <DAW>

A:Cross-references: EMBL:X76907; NID:g439525; PID:g439526

C:Superfamily: gamma-casein

Query Match

Best Local Similarity 81.1%; Score 30; DB 2; Length 180;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIK 7

|||||:

Db 79 KQLKRIK 85

RESULT 14

D69340

cobalamin biosynthesis precorrin-3 methylase (cbiH) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1998

C:Accession: D69340

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A:Reference number: A69250; MUID:98049343

A:Accession: D69340

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-446 <KLE>

A:Cross-references: GB:AE001055; GB:AE000782; NID:g2689378; PID:g2649892; TIGR:AF0724

A:Note: Region (350-\*) is homologous to other bacterial precorrin isomerases; Acc B69

C:Superfamily: probable cobalamin biosynthesis precorrin-3 methylase

Query Match

Best Local Similarity 81.1%; Score 30; DB 2; Length 446;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELKRIKI 8

|||||:

Db 63 KELDRVKI 70

RESULT 15

S53058

Probable membrane protein YMR128w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YMR553.04

C:Species: Saccharomyces cerevisiae

C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 05-Jun-1998

C:Accession: S53058

R:Badcock, K.; Churcher, C.

submitted to the EMBL Data Library, March 1995

A:Reference number: S53055

A:Accession: S53058

A:Molecule type: DNA

A:Residues: 1-1267 <BAD>

A:Cross-references: EMBL:Z48622; NID:g728663; PID:g728667; MIPS:YMR128w

C:Genetics:

A:Gene: SGD:ECM16

A:Cross-references: SGD:S0004735; MIPS:YMR128w

```
Query Match      83.8%; Score 31; DB 2; Length 309;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8
   |::|::|
Db 127 QQLKKIKI 134

RESULT 7
C70226
hypothetical protein BBE16 - Lyme disease spirochete plasmid E/lp25
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 05-Jun-1998
C:Accession: C70226
R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: C70226
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-206 <KLE>
A:Cross-references: GB:AE000785; NID:g2689951; PID:g2689979; TIGR:BBE16
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match      83.8%; Score 31; DB 2; Length 206;
Best Local Similarity 75.0%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8
   |::|::|
Db 176 KKLKRLKI 183

RESULT 8
C64440
DNA repair protein RAD24 homolog (intein-containing) - Methanococcus jannaschii
N:Contains: intein
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Jul-1998
C:Accession: C64440
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1038-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999
A:Accession: C64440
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1195 <BUL>
A:Cross-references: GB:067555; GB:L77117; NID:g1591760; PID:g1499976; TIGR:MJ1124
C:Genetics:
A:Superfamily: DEAD/H box helicase homology
C:Keywords: protein splicing
F:39-372/Domain: DEAD/H_box helicase homology <DEAD>

Query Match      83.8%; Score 31; DB 2; Length 1195;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8
   |::|::|
```

```
Db 802 KEIKIKV 809

RESULT 9
S67760
hypothetical protein YDL201w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D1075
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 14-Nov-1997
C:Accession: S67760
R:Schmidt, E.R.; Bahr, A.; Kraemer, C.; Hankeln, T.; Moeller-Rieker, S.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67756
A:Accession: S67760
A:Molecule type: DNA
A:Residues: 1-286 <SCH>
A:Cross-references: EMBL:Z74249; NID:g1431331; PID:e253107; PID:g1431332; MIPS:YDL201
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 4L

Query Match      83.8%; Score 31; DB 2; Length 286;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELKRIKI 8
   |::|::|
Db 26 KELKHVKI 33

RESULT 10
E71606
hypothetical protein PFB0765w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: E71606
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: E71606
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-980 <GAR>
A:Cross-references: GB:AE001417; GB:AE001362; NID:g3845271; PID:g3845272; TIGR:PFB076
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0765w

Query Match      83.8%; Score 31; DB 2; Length 980;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIK 7
   |::|::|
Db 570 KELARLK 576

RESULT 11
RNZA2
genome polyprotein - human respiratory syncytial virus (strain A2)
N:Alternate names: polymerase L protein
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: human respiratory syncytial virus
A:Note: host Homo sapiens (man)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C:Accession: A40317; A28319; P80048
R:Stec, D.S.; Hill III, M.G.; Collins, P.L.
Virology 183, 273-287, 1991
A:Title: Sequence analysis of the polymerase L gene of human respiratory syncytial vi
```

Query Match 100.0%; Score 37; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
|||||  
Db 58 KELKRIKI 65

RESULT 2  
S43383  
bactericidal/permeability-increasing protein - synthetic  
C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in *Escherichia coli*  
C:Date: 20-Oct-1994 #sequence\_revision 15-Feb-1996 #text\_change 15-Feb-1996  
R:Q1, S.Y.; Li, Y.; O'Connor, C.D.  
Biochem. J. 258, 711-718, 1994  
A:Title: The region around residue 115 of human bactericidal/permeability-increasing protein of a gene coding for the active domain and characterization of recombinant proteins.  
A:Reference number: S43383  
A:Accession: S43383  
A:Molecule type: DNA  
A:Residues: 1-250 <Q1S>

Query Match 100.0%; Score 37; DB 4; Length 250;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
|||||  
Db 28 KELKRIKI 35

RESULT 3  
B71162  
hypothetical protein PH0495 - *Pyrococcus horikoshii*  
C:Species: *Pyrococcus horikoshii*  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 14-Aug-1998  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.A.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A:Reference number: A71000; MUID:98344137  
A:Accession: B71162  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-440 <KAW>  
A:Cross-references: GB:AP000002; NID:g3236129; PID:d1030526; PID:g3256900  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0495

Query Match 97.3%; Score 36; DB 2; Length 440;  
Best Local Similarity 87.5%; Pred. No. 5.5;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
|||||  
Db 86 KELKRIKV 93

RESULT 4  
F70361  
tRNA pseudouridine 55 synthase - Aquifex aeolicus  
C:Species: *Aquifex aeolicus*  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 21-Aug-1998  
A:Accession: F70361  
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy

V. Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
A:Reference number: A70300; MUID:98196666  
A:Accession: F70361  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-287 <AQF>  
A:Cross-references: GB:AE000703; NID:g2983287; PID:g2983293; GB:AE000657  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: truD  
C:Superfamily: *Escherichia coli* protein P35

Query Match 86.5%; Score 32; DB 2; Length 287;  
Best Local Similarity 75.0%; Pred. No. 23;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
|||||  
Db 190 KELRLKI 197

RESULT 5  
A47447  
calcium channel protein alpha-1 chain (variant doe-1) - electric ray (*Discopyge ommat*  
C:Species: *Discopyge ommat*  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 03-Mar-1995  
C:Accession: A47447  
R:Horne, W.A.; Ellinor, P.T.; Imman, I.; Zhou, M.; Tsien, R.W.; Schwarz, T.L.  
Proc. Natl. Acad. Sci. U.S.A. 90, 3787-3791, 1993  
A:Title: Molecular diversity of Ca(2+) channel alpha 1 subunits from the marine ray D  
A:Reference number: A47447; MUID:93248175  
A:Accession: A47447  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-2223 <HOR>  
A:Note: sequence extracted from NCBI backbone (NCBIP:130671)

Query Match 86.5%; Score 32; DB 2; Length 2223;  
Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
|||||  
Db 1904 KELKRIKL 1911

RESULT 6  
F70217  
conserved hypothetical protein BBB10 - Lyme disease spirochete plasmid B/cp26  
C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
C:Accession: F70217  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.  
A:Reference number: A70100; MUID:98065943  
A:Accession: F70217  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-309 <KLE>  
A:Cross-references: GB:AE000792; NID:g2689884; PID:g2689896; TIGR:BBB10  
A:Experimental source: strain B31  
C:Genetics:  
A:Genome: plasmid

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:38 : Search time 105.14 Seconds  
(without alignments)  
3.049 Million cell updates/sec

Title: US-09-124-280A-25

Perfect score: 37  
Sequence: 1 KELKRIKI 8

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	487	2	A30909 bactericidal/perme
2	37	100.0	250	4	S43383 bactericidal/perme
3	36	97.3	440	2	B71162 hypothetical prote
4	32	86.5	287	2	F70361 tRNA pseudouridine
5	32	86.5	2223	2	A47447 calcium channel pr
6	31	83.8	309	2	F70217 conserved hypothet
7	31	83.8	206	2	C70226 hypothetical prote
8	31	83.8	1195	2	C64440 DNA repair protein
9	31	83.8	286	2	S67760 hypothetical prote
10	31	83.8	980	2	E71606 genome polypeptide
11	30	81.1	2165	1	RNN2A2 epidermal growth f
12	30	81.1	1308	2	A47253 alpha-s2-casein fo
13	30	81.1	180	2	S39775 cobalamin biosynth
14	30	81.1	446	2	D69340 probable membrane
15	30	81.1	1267	2	S33058 conserved hypothet
16	30	81.1	218	2	H64638 hypothetical prote
17	30	81.1	103	2	G64576 hypothetical prote
18	30	81.1	245	2	E69310 CAMP-dependent cel
19	30	81.1	1148	2	JC5984 DNA (cytosine-5)-
20	30	81.1	1490	2	JC5145 microfilament-associ
21	30	81.1	442	2	A42670 microfilament-associ
22	30	81.1	439	2	A35565 microfilament-associ
23	30	81.1	788	2	A71076 hypothetical prote
24	30	81.1	67	2	F71150 hypothetical prote
25	30	81.1	499	2	C71111 probable phenylala
26	29	78.4	1830	1	S19188 myosin-V - chicken
27	29	78.4	4563	1	LPHUB apolipoprotein B-1
28	29	78.4	373	2	G70355 8-amino-7-oxononan
29	29	78.4	812	2	A53016 myosin heavy chain
30	29	78.4	1156	2	B70356 chromosome assembl
31	29	78.4	186	2	S73141 hypothetical prote
32	29	78.4	449	2	S27098 pectate lyase (EC
33	29	78.4	208	2	A70253 antigen, P35 homol
34	29	78.4	114	2	S32226 hypothetical prote
35	29	78.4	1008	2	F71727 acriflavin resista
36	29	78.4	241	2	F71478 probable pseudouri
37	29	78.4	211	2	C64336 hypothetical prote
38	29	78.4	486	2	E64438 hypothetical prote
39	29	78.4	4572	2	S57908 hypothetical polyp

#### ALIGNMENTS

##### RESULT 1

A30909  
bactericidal/permeability-increasing protein precursor - human  
N:Alternate names: 55K bactericidal protein  
C:Species: Homo sapiens (man)  
C:Date: 18-Apr-1989 #sequence, revision 18-Apr-1989 #text, change 17-Mar-1999  
C:Accession: A33850; B54136; A29464; A43600; A49716; A30909  
R:Gray, P.W.; Flagg, G.; Leong, S.R.; Gumina, R.J.; Weiss, J.; Ooi, C.E.; Elsbach, P.  
J. Biol. Chem. 264, 9505-9509, 1989  
A:Title: Cloning of the cDNA of a human neutrophil bactericidal protein. Structural a  
A:Reference number: A33850; MUID:89255455  
A:Accession: A33850  
A:Molecule type: mRNA  
A:Residues: 1-487 <GRA>  
A:Cross-references: GB:J04739; NID:gl79528; PID:gl79529  
R:Wilde, C.G.; Seilhamer, J.J.; McGrogan, M.; Ashton, N.; Snable, J.L.; Lane, J.C.; L  
J. Biol. Chem. 269, 17411-17416, 1994  
A:Title: Bactericidal/permeability-increasing protein and lipopolysaccharide (LPS)-bi  
A:Reference number: A54136; MUID:94292492  
A:Accession: B54136  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-374, 'L', 376-487 <WIL>  
A:Experimental source: HL-60 cells  
A:Note: sequence extracted from NCBI backbone (NCBIP:149855)  
R:Ooi, C.E.; Weiss, J.; Elsbach, P.; Frangione, B.; Mannion, B.  
J. Biol. Chem. 262, 14891-14894, 1987  
A:Title: A 25-kDa amino-terminal fragment carries all the antibacterial activities of  
A:Reference number: A29464; MUID:88033057  
A:Accession: A29464  
A:Molecule type: protein  
A:Residues: 32-51 <OOI>  
A:Experimental source: neutrophils  
R:Wasiluk, K.R.; Skubitz, K.M.; Gray, B.H.  
Infect. Immun. 59, 4193-4200, 1991  
A:Title: Comparison of granule proteins from human polymorphonuclear leukocytes which  
A:Reference number: A43600; MUID:92040097  
A:Accession: A43600  
A:Molecule type: protein  
A:Residues: 32-52, 'R' <WAS>  
R:Little, R.G.; Keiner, D.N.; Lim, E.; Burke, D.J.; Conlon, P.J.  
J. Biol. Chem. 269, 1865-1872, 1994  
A:Title: Functional domains of recombinant bactericidal/permeability increasing prote  
A:Reference number: A49716; MUID:94124531  
A:Accession: A49716  
A:Molecule type: protein  
A:Residues: 32-130, 132-141, 143-165, 202-215, 'E', 217-225 <LIT>  
C:Comment: The bactericidal/permeability-increasing protein (BPI) is a 60 kD membrane  
which is specific for gram-negative bacteria. BPI has a high affinity for lipopolysac  
between BPI and an LPS-binding protein from liver and cholesterol ester transfer prot  
C:Genetics:  
A:Gene: GDB:BPI  
A:Cross-references: GDB:131572; OMIM:109195  
A:Map position: 20q11.23-20q12  
C:Superfamily: lipopolysaccharide-binding protein  
C:Keywords: antibacterial; cytotoxin; glycoprotein; heparin binding; neutrophil  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-487/Product: bactericidal permeability-increasing protein #status predicted <WAT  
F:32-51/Region: bactericidal #status predicted  
F:380/Binding site: carbohydrate (Asn) (covalent) #status predicted

hypothetical prote  
phosphatidylserine  
poly(A)-specific r  
hypothetical prote  
hypothetical prote  
3beta-hydroxy-Delt

40 29 78.4 427 2 S38002  
41 29 78.4 1138 2 S64484  
42 29 78.4 1480 2 S48440  
43 29 78.4 858 2 T00258  
44 29 78.4 361 2 C71242  
45 28 75.7 373 1 DEHUHS



**THIS PAGE BLANK (US)**

NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 27129/32043  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-232-527-2

MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "Domain I"  
US-08-372-783-1

Query Match 100.0%; Score 37; DB 1; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.41;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
      |||||||  
Db 11 KELKRIKI 18

Search completed: September 7, 1999, 23:07:46  
Job time: 1744 sec

Query Match 100.0%; Score 37; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
      |||||||  
Db 58 KELKRIKI 65

RESULT 15  
US-08-372-783-1  
; Sequence 1, Application US/08372783  
; Patent No. 5578572  
; GENERAL INFORMATION:  
; APPLICANT: Horwitz, Arnold H.  
; APPLICANT: Lambert, Lewis H.  
; APPLICANT: Little, Roger G.  
; TITLE OF INVENTION: Anti-Gram-Positive Bacterial Methods and  
; TITLE OF INVENTION: Materials  
; NUMBER OF SEQUENCES: 237  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/372,783  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,540  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,762  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,222  
; FILING DATE: 14-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27129/32415  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 29 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

```
; APPLICANT: Elsbach, Peter
; APPLICANT: Weiss, Jerrold
; TITLE OF INVENTION: Biologically Active
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing Protein Fragments
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howard M. Frankfort
; STREET: 805 Third Ave.
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,968
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/754,204
; FILING DATE: 26-AUG-1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Frankfort, Howard M.
; REGISTRATION NUMBER: 32,613
; REFERENCE/DOCKET NUMBER: 5969/059820US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: blood
; CELL TYPE: promyelocytic leukemia cells
; US-08-173-968-2
```

```
Query Match 100.0%; Score 37; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. NO. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 KELKRIKI 8
Db 58 KELKRIKI 65
```

```
RESULT 13
US-08-173-968-3
; Sequence 3, Application US/08173968
; Patent No. 5576292
; GENERAL INFORMATION:
; APPLICANT: Elsbach, Peter
; APPLICANT: Weiss, Jerrold
; TITLE OF INVENTION: Biologically Active
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing Protein Fragments
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howard M. Frankfort
; STREET: 805 Third Ave.
; CITY: New York
; STATE: NY
; COUNTRY: USA
```

```
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,968
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/754,204
; FILING DATE: 26-AUG-1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Frankfort, Howard M.
; REGISTRATION NUMBER: 32,613
; REFERENCE/DOCKET NUMBER: 5969/059820US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: blood
; CELL LINE: promyelocytic leukemia cells
; US-08-173-968-3
```

```
Query Match 100.0%; Score 37; DB 1; Length 221;
Best Local Similarity 100.0%; Pred. NO. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 KELKRIKI 8
Db 27 KELKRIKI 34
```

```
RESULT 14
US-08-232-527-2
; Sequence 2, Application US/08232527
; Patent No. 5578568
; GENERAL INFORMATION:
; APPLICANT: Ammons, William Steve et al.
; TITLE OF INVENTION: Method of Treating Conditions Associated with
; TITLE OF INVENTION: Intestinal Ischemia/Reperfusion
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,527
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
```

us-09-124-280a-25.ra1

.Wed Sep 8 09:43:26 1999

Best Local Similarity 100.0%; Pred. No. 5.2; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;

;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "BPI.54"  
US-08-311-611A-5

Query Match 100.0%; Score 37; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.22; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;

QY 1 KELKRIKI 8  
Db 7 KELKRIKI 14

RESULT 10  
US-08-311-611A-69  
; Sequence 69, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "BPI"  
US-08-311-611A-69

Query Match 100.0%; Score 37; DB 1; Length 487;

QY 1 KELKRIKI 8  
Db 58 KELKRIKI 65

RESULT 11  
US-08-311-611A-146  
; Sequence 146, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 146:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-311-611A-146

Query Match 100.0%; Score 37; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 5.2; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;

QY 1 KELKRIKI 8  
Db 58 KELKRIKI 65

RESULT 12  
US-08-173-968-2  
; Sequence 2, Application US/08173968  
; Patent No. 5576292  
; GENERAL INFORMATION:

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 30 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "BPI.14"  
US-08-311-611A-2

Query Match 100.0%; Score 37; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. NO. 0.42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
DB 7 KELKRIKI 14

RESULT 8  
US-08-311-611A-3  
; Sequence 3, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:

; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "BPI.4"  
US-08-311-611A-3

Query Match 100.0%; Score 37; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
DB 3 KELKRIKI 10

RESULT 9  
US-08-311-611A-5  
; Sequence 5, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids

us-09-124-280a-25.ra1

Wed Sep 8 09:43:26 1999

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/311.611A

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/273.401

FILING DATE: 11-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/125.651

FILING DATE: 22-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Sharp, Jeffrey S.

REGISTRATION NUMBER: 31,879

REFERENCE/DOCKET NUMBER: 32251

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: "Domain I"

US-08-311-611A-1

Query Match

Best Local Similarity 100.0%; Score 37; DB 1; Length 29;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8

Db 11 KELKRIKI 18

RESULT 7

US-08-311-611A-2

Sequence 2, Application US/08311611A

Patent No. 5523288

GENERAL INFORMATION:

APPLICANT: Cohen, Jonathan

APPLICANT: Kung, Ada H.C.

TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial

TITLE OF INVENTION: Infection by Administration of

TITLE OF INVENTION: Bactericidal/Permeability-Increasing

TITLE OF INVENTION:

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &amp; Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

QY 1 KELKRIKI 8

Db 58 KELKRIKI 65

RESULT 5

US-08-414-924-2

Sequence 2, Application US/08414924

Patent No. 5494896

GENERAL INFORMATION:

APPLICANT: Hansbrough, John F.

TITLE OF INVENTION: Method of Treating Conditions

TITLE OF INVENTION: Associated With Burn Injuries

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/414.924

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sharp, Jeffrey S.

REGISTRATION NUMBER: 31,879

REFERENCE/DOCKET NUMBER: 27129/32297

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 487 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-414-924-2

Query Match 100.0%; Score 37; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 5.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8

Db 58 KELKRIKI 65

RESULT 6

US-08-311-611A-1

Sequence 1, Application US/08311611A

Patent No. 5523288

GENERAL INFORMATION:

APPLICANT: Cohen, Jonathan

APPLICANT: Kung, Ada H.C.

TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial

TITLE OF INVENTION: Infection by Administration of

TITLE OF INVENTION: Bactericidal/Permeability-Increasing

TITLE OF INVENTION:

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &amp; Borun

; APPLICANT: Grinna, Lynn S  
; TITLE OF INVENTION: Stable Bactericidal/Permeability-  
; TITLE OF INVENTION: Increasing Protein Products and Pharmaceutical  
; TITLE OF INVENTION: Compositions Containing the Same  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/013.801  
; FILING DATE: 02 FEB 1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Thomas C.  
; REGISTRATION NUMBER: P-36,989  
; REFERENCE/DOCKET NUMBER: 27129/30911  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/346-5750  
; TELEFAX: 312/346-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-013-801-2

Query Match 100.0%; Score 37; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
Db 58 KELKRIKI 65

RESULT 3  
US-08-072-063-2  
; Sequence 2, Application US/08072063  
; Patent No. 5439807  
; GENERAL INFORMATION:  
; APPLICANT: Theofan, Georgia  
; APPLICANT: Grinna, Lynn S  
; APPLICANT: Horwitz, Arnold  
; TITLE OF INVENTION: BPI-Immunoglobulin Fusion Proteins  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/072,063  
; FILING DATE: 19930519

; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers Thomas C.  
; REGISTRATION NUMBER: 36,989  
; REFERENCE/DOCKET NUMBER: 30659  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-072-063-2

Query Match 100.0%; Score 37; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
Db 58 KELKRIKI 65

RESULT 4  
US-08-212-132-2  
; Sequence 2, Application US/08212132  
; Patent No. 5447913  
; GENERAL INFORMATION:  
; APPLICANT: Little, Roger G.  
; APPLICANT: Ammons, William Steve  
; TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-  
; TITLE OF INVENTION: Increasing Protein Dimer Products  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/212,132  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 27129/31735  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-212-132-2

Query Match 100.0%; Score 37; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:46 ; Search time 82.37 Seconds  
(without alignments)  
0.958 Million cell updates/sec

Title: US-09-124-280A-25  
Perfect score: 37  
Sequence: 1 KELKRIKI 8

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	487	1	US-08-030-644-2
2	37	100.0	487	1	US-08-013-801-2
3	37	100.0	487	1	US-08-072-063-2
4	37	100.0	487	1	US-08-212-132-2
5	37	100.0	487	1	US-08-414-924-2
6	37	100.0	29	1	US-08-311-611A-1
7	37	100.0	30	1	US-08-311-611A-2
8	37	100.0	22	1	US-08-311-611A-3
9	37	100.0	15	1	US-08-311-611A-5
10	37	100.0	487	1	US-08-311-611A-69
11	37	100.0	487	1	US-08-311-611A-146
12	37	100.0	487	1	US-08-173-968-2
13	37	100.0	221	1	US-08-232-527-2
14	37	100.0	487	1	US-08-372-783-1
15	37	100.0	29	1	US-08-372-783-2
16	37	100.0	30	1	US-08-372-783-3
17	37	100.0	22	1	US-08-372-783-5
18	37	100.0	15	1	US-08-372-783-5
19	37	100.0	487	1	US-08-372-783-69
20	37	100.0	487	1	US-08-372-783-146
21	37	100.0	29	1	US-08-372-105-1
22	37	100.0	30	1	US-08-372-105-2
23	37	100.0	22	1	US-08-372-105-3
24	37	100.0	15	1	US-08-372-105-5
25	37	100.0	487	1	US-08-372-105-69
26	37	100.0	487	1	US-08-372-105-146
27	37	100.0	487	1	US-08-415-158-2
28	37	100.0	487	1	US-08-064-693-2
29	37	100.0	487	1	US-08-291-112-2
30	37	100.0	8	1	US-08-097-830E-25
31	37	100.0	29	1	US-08-306-473A-1
32	37	100.0	30	1	US-08-306-473A-2
33	37	100.0	22	1	US-08-306-473A-3
34	37	100.0	15	1	US-08-306-473A-5
35	37	100.0	487	1	US-08-306-473A-69
36	37	100.0	487	1	US-08-306-473A-146
37	37	100.0	487	1	US-08-430-417-2
38	37	100.0	487	1	US-08-557-287-2
39	37	100.0	487	1	US-08-470-366-2

ALIGNMENTS

RESULT 1

US-08-030-644-2

; Sequence 2, Application US/08030644

; Patent No. 5348942

; GENERAL INFORMATION:

; APPLICANT: Little, Roger G. II

; APPLICANT: Gazzano-Santoro, Helene

; APPLICANT: Parent, James Brian

; TITLE OF INVENTION: Therapeutic uses of

; Bactericidal/Permeability-Increasing Protein Products

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/030,644

; FILING DATE: 19930312

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Sharp, Jeffrey S.

; REGISTRATION NUMBER: 31,879

; REFERENCE/DOCKET NUMBER: 31229

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300

; TELEFAX: 312/474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 487 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-030-644-2

Query Match 100.0%; Score 37; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 5-2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8

|||||

Db 58 KELKRIKI 65

RESULT 2

US-08-013-801-2

; Sequence 2, Application US/08013801

; Patent No. 5420019

; GENERAL INFORMATION:

; APPLICANT: Theofan, Georgia

; APPLICANT: Horwitz, Arnold

; APPLICANT: Burke, David

; APPLICANT: Baltanian, Manik



Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KELKRIKI 8
Db 7 KELKRIKI 14

RESULT 14
R63684
ID R63684 standard; peptide: 22 AA.
AC R63684;
DT 01-MAY-1995 (first entry)
DE BPI domain I.
KW Human; bactericidal/permeability-increasing protein; BPI; heparin;
KW binding agent; neutralisation; anti-coagulant effect; inhibition;
KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;
KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;
KW lipopolysaccharide; circulation; lymphokine; compromised immune response; microbe;
KW macrophage; activation; lymphokine; decontaminating; Helicobacter;
KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;
KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.
OS Homo sapiens.
PN WO9420532-A.
PD 15-SEP-1994.
PF 11-MAR-1994; U02465.
PR 12-MAR-1993; US-030644.
PR 15-JUL-1993; US-093202.
PR 14-JAN-1994; US-183222.
PA (XOMA ) XOMA CORP.
PI Little RG;
DR WPI: 94-302964/37.
PT New human bactericidal permeability increasing peptides - derived
PT from the functional domains of BPI and having BPI activities such
PT as bactericidal activity
PS Claim 2; Page 157; 254pp; English.
CC The sequences given in R63682-750, R62087-100 and R62491-500 are
CC peptides derived from human bactericidal/permeability-increasing
CC protein (BPI). The sequences given in R63682-86 are derived from
CC positions 17-45 of BPI. Peptides such as these may be used as
CC heparin binding agents, for neutralising the anti-coagulant effect
CC of heparin, for inhibiting angiogenesis, eg. associated with ocular
CC retinopathy, for inhibiting endothelial cell proliferation, for
CC contraception, for inhibiting malignant tumour cell proliferation,
CC for treating a chronic inflammatory disease state, eg. rheumatoid
CC arthritis, and for treating gram-negative bacterial infection. The
CC peptides may also be used for treating a subject suffering from the
CC adverse effects of the presence of lipopolysaccharide in the circulation,
CC eg. a compromised immune response to microbes or tumour cells due to
CC inhibition of macrophage activation by T-cell lymphokines or increased
CC production of a cytokine, for decontaminating a fluid containing
CC lipopolysaccharide or for treating a disease associated with Helicobacter
CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.
CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B
CC or cefamandole nafate or LBP protein products. The peptides are pref.
CC prepared by solid phase synthesis.
SQ Sequence 22 AA;

```

Query Match 100.0%; Score 37; DB 1; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 0.36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KELKRIKI 8
Db 3 KELKRIKI 10

RESULT 15
R63682
ID R63682 standard; peptide: 29 AA.
AC R63682;
DT 01-MAY-1995 (first entry)

```

```

DE BPI domain I.
KW Human; bactericidal/permeability-increasing protein; BPI; heparin;
KW binding agent; neutralisation; anti-coagulant effect; inhibition;
KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;
KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;
KW lipopolysaccharide; circulation; lymphokine; compromised immune response; microbe;
KW macrophage; activation; lymphokine; decontaminating; Helicobacter;
KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;
KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.
OS Homo sapiens.
PN WO9420532-A.
PD 15-SEP-1994.
PF 11-MAR-1994; U02465.
PR 12-MAR-1993; US-030644.
PR 15-JUL-1993; US-093202.
PR 14-JAN-1994; US-183222.
PA (XOMA ) XOMA CORP.
PI Little RG;
DR WPI: 94-302964/37.
PT New human bactericidal permeability increasing peptides - derived
PT from the functional domains of BPI and having BPI activities such
PT as bactericidal activity
PS Claim 2; Page 157; 254pp; English.
CC The sequences given in R63682-750, R62087-100 and R62491-500 are
CC peptides derived from human bactericidal/permeability-increasing
CC protein (BPI). The sequences given in R63682-86 are derived from
CC positions 17-45 of BPI. Peptides such as these may be used as
CC heparin binding agents, for neutralising the anti-coagulant effect
CC of heparin, for inhibiting angiogenesis, eg. associated with ocular
CC retinopathy, for inhibiting endothelial cell proliferation, for
CC contraception, for inhibiting malignant tumour cell proliferation,
CC for treating a chronic inflammatory disease state, eg. rheumatoid
CC arthritis, and for treating gram-negative bacterial infection. The
CC peptides may also be used for treating a subject suffering from the
CC adverse effects of the presence of lipopolysaccharide in the circulation,
CC eg. a compromised immune response to microbes or tumour cells due to
CC inhibition of macrophage activation by T-cell lymphokines or increased
CC production of a cytokine, for decontaminating a fluid containing
CC lipopolysaccharide or for treating a disease associated with Helicobacter
CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.
CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B
CC or cefamandole nafate or LBP protein products. The peptides are pref.
CC prepared by solid phase synthesis.
SQ Sequence 29 AA;

```

Query Match 100.0%; Score 37; DB 1; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.47;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KELKRIKI 8
Db 11 KELKRIKI 18

```

Search completed: September 7, 1999, 22:49:57  
 Job time: 7878 sec

PA (XOMA ) XOMA CORP  
 PI Baltian M, Burke D, Grinna L, Horwitz A, Theofan G;  
 DR WPI: 94-279744/34.  
 DR N-PSDB: 067270.  
 PT Bactericidal-permeability-increasing protein analogs with  
 PT Cys132 or Cys135 replaced - also fusion proteins and C-terminally  
 PT truncated forms, have increased stability and are useful to treat  
 PT gram-negative bacterial infection  
 PS Disclosure: Page 52-54; 77pp; English.  
 CC Analogs, fusion proteins and C-terminal truncated forms of human BPI  
 CC are claimed that show improved resistance to dimerization and  
 CC adduct formation.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 37; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 Db 58 KELKRIKI 65  
 |||||

RESULT 12  
 R62370 standard; Protein: 487 AA.  
 AC R62370;  
 DE Recombinant bactericidal/permeability-increasing protein (rBPI-50).  
 KW N-terminal; recombinant; bactericidal/permeability protein; rBPI;  
 KW signal protein; human; rBPI-23; binding assay; heparin-binding; domain;  
 KW Limulus amoebocyte lysate; LAL; inhibition; bactericidal activity;  
 KW chronic inflammatory disease; arthritis; psoriasis; Crohn's disease;  
 KW inflammatory bowel disease; asthma.  
 OS Homo sapiens.  
 FH Key  
 FT Peptide  
 FT 1..31 Location/Qualifiers  
 FT /note= "Signal peptide"  
 FT 32..487  
 FT /note= "mature protein"  
 FT 17..55  
 FT /note= "LAL inhibiting domain"  
 FT 21..55  
 FT /note= "Heparin binding site"  
 FT 65..107  
 FT /note= "Heparin binding site"  
 FT 73..99  
 FT /note= "LAL inhibiting domain"  
 FT 85..99  
 FT /note= "Bactericidally active fragment"  
 FT 137..163  
 FT /note= "LAL inhibiting domain"  
 FT 137..171  
 FT /note= "Heparin binding site"  
 FT W09420128-A.  
 PN 15-SEP-1994.  
 PD 11-MAR-1994; U02401.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PA (XOMA ) XOMA CORP.  
 PI Gazzano-Santoro H, Little RG, Parent JB;  
 DR N-PSDB: Q73382.  
 DR Uses of bactericidal permeability increasing protein product - to  
 PT neutralise anticoagulant effect of heparin, to inhibit  
 PT angiogenesis or cell proliferation, as contraceptive and to reduce  
 PT inflammation  
 PS Disclosure: Page 43-45; 78pp; English.  
 CC This sequence represents the N-terminal of recombinant bactericidal/  
 CC permeability (rBPI) protein. This expression product encoding the 31  
 CC residue signal protein and the first 199 amino acids of the N-terminal  
 CC of mature human BPI has a molecular weight of 23 kD and is designated

CC rBPI-23. In binding assays rBPI-23 was shown to have three separate  
 CC functional domains with heparin-binding activity, at least three major  
 CC domains with significant Limulus amoebocyte lysate (LAL) inhibition and  
 CC one region with bactericidal activity. BPI, or active fragments of it  
 CC may be used to treat and prevent the effects of chronic inflammatory  
 CC disease states such as arthritis, psoriasis, inflammatory bowel disease,  
 CC Crohn's disease and asthma.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 37; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 Db 58 KELKRIKI 65  
 |||||

RESULT 13  
 R63686 standard; peptide: 15 AA.  
 ID R63686;  
 AC R63686;  
 DT 01-MAY-1995 (first entry)  
 DE BPI derived peptide, BPI-54.  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW contraception; malignant; tumour cell; inflammatory disease; T-cell;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipopolysaccharide; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 PN W09420532-A.  
 PD 15-SEP-1994.  
 PD 11-MAR-1994; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.  
 PA (XOMA ) XOMA CORP.  
 PI Little RG;  
 DR WPI: 94-302964/37.  
 FT New human bactericidal permeability increasing peptides - derived  
 FT from the functional domains of BPI and having BPI activities such  
 FT as bactericidal activity  
 PS Claim 2; Page 157; 254pp; English.  
 CC The sequences given in R63682-750, R62087-100 and R62491-500 are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC positions 17-45 of BPI. Peptides such as these may be used as  
 CC protein (BPI). The sequences given in R63682-86 are derived from  
 CC heparin binding agents, for neutralising the anti-coagulant effect  
 CC of heparin, for inhibiting angiogenesis, eg. associated with ocular  
 CC retinopathy, for inhibiting endothelial cell proliferation, for  
 CC contraception, for inhibiting malignant tumour cell proliferation,  
 CC for treating a chronic inflammatory disease state, eg. rheumatoid  
 CC arthritis, and for treating gram-negative bacterial infection. The  
 CC peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipopolysaccharide in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipopolysaccharide or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 SQ Sequence 15 AA;

Query Match 100.0%; Score 37; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.25;

PT active form  
 PS Disclosure; Fig 7; 53pp; English.  
 CC The human bactericidal/permeability increasing protein may be used  
 CC in a compsn. solubilised in a lipid carrier. The BPIP binds  
 CC lipopolysaccharides and inhibits the immunostimulatory and toxic  
 CC activities of LPS in vitro and in vivo. The lipid compsn. stably  
 CC maintains the BPIP in a sol. active form. The compsn. can be used  
 CC for detecting or quantifying LPS in a sample, coating a surgical  
 CC tool, coating an implantable invasive device, decontaminating a  
 CC fluid contg. LPC and treating endotoxin related -shock.  
 CC -disseminated intravascular coagulation, -anaemia, -thrombocytopenia,  
 CC -adult respiratory stress syndrome or -renal failure, or for  
 CC preventing endotoxaemia.  
 CC See also R33776, R33729-31.  
 SQ Sequence 243 AA;

Query Match 100.0%; Score 37; DB 1; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 3.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KELKRIKI 8  
 |||||  
 Db 58 KELKRIKI 65

RESULT 9  
 ID R44420 standard; Protein; 487 AA.  
 AC R44420; 25-NOV-1993.  
 DT 03-JUN-1994 (first entry)  
 DE PING4512 encoded fusion protein.  
 KW Polymerase chain reaction; primer: amplify; PCR; plasmid; PING4512;  
 KW PING4503; bactericidal/permeability-increasing protein; BPI; IgG;  
 KW heavy; chain; pMB27; fusion protein; Gram negative; renal failure;  
 KW bacterial infection; endotoxin related shock; metabolic acidosis;  
 KW disseminated intravascular coagulation; anaemia; thrombocytopenia;  
 KW leukopenia; adult respiratory distress syndrome; hypotension; fever;  
 KW constant region; lipopolysaccharide complement binding;  
 KW placentar transfer; Fc receptor binding.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT peptide 1..31  
 FT protein /note= "Signal peptide"  
 FT protein 32..487  
 FT protein /note= "Mature rBPI(191)-IgG fusion protein"

PN W09323434-A.  
 PD 25-NOV-1993.  
 PF 19-MAY-1993; U04754.  
 PR 19-MAY-1992; US-885911.  
 PA (XOMA ) XOMA CORP.  
 PI Grinna LS, Horwitz A, Theofan G;  
 DR WPI: 93-386485/48.  
 DR N-PSDB: Q52488.  
 PT New fusion proteins for treating bacterial infections - comprising  
 PT a bactericidal-permeability-increasing protein and an immunoglobulin  
 PT heavy chain constant domain  
 PS Claim 20: Page 48-50: 75pp; English.  
 CC This sequence is encoded by the plasmid PING4512. PING4512 was  
 CC constructed using the primer sequence given in Q52480. Plasmid  
 CC PING4503 was cut with AlwNI, the ends made blunt with T4 DNA  
 CC polymerase, and the DNA then cut with SalI. The resultant approx.  
 CC 700bp SalI/blunt fragment contains the 30 bp 5' untranslated region  
 CC and the DNA encoding the signal peptide and first 191 amino acids of  
 CC bactericidal/permeability-increasing protein (BPI). IgG heavy chain  
 CC (HC) sequences, which also included the hinge region of the IgG HC  
 CC were PCR amplified from plasmid pMB27 using these primers. Plasmids  
 CC such as PING4512 may be used to produce fusion proteins which are  
 CC useful for the treatment of Gram negative bacterial infections and  
 CC their sequelae including endotoxin related shock and conditions  
 CC associated with it, such as disseminated intravascular coagulation,  
 CC anemia, thrombocytopenia, leukopenia, adult respiratory distress  
 CC syndrome, renal failure, hypotension, fever and metabolic acidosis.

CC Proving BPI or a fragment of it, as part of the fusion with an  
 CC immunoglobulin heavy chain constant region provides the potential  
 CC advantages of Fc receptor binding, bivalent binding to lipopoly-  
 CC saccharide complement binding and increasing placental transfer.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 37; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KELKRIKI 8  
 |||||  
 Db 58 KELKRIKI 65

RESULT 10  
 R43886  
 ID R43886 standard; Protein; 487 AA.  
 AC R43886;  
 DT 01-JUN-1994 (first entry)  
 DE Human bactericidal/permeability-increasing protein.  
 KW Bactericidal/permeability-increasing protein;  
 KW endotoxin binding protein.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT protein 32..487  
 FT protein /label= human BPI protein  
 FT misc\_difference 1..31  
 FT /label= signal peptide  
 PN W09323540-A.  
 PD 25-NOV-1993.  
 PF 19-MAY-1993; U04752.  
 PR 19-MAY-1992; US-885501.  
 PA (XOMA ) XOMA CORP.  
 PI Grinna LS;  
 DR WPI: 93-386572/48.  
 DR P-PSDB: Q52268.  
 PT Recovering endotoxin binding protein from transformed cell  
 PT cultures - by retention on cation exchange material and elution,  
 PT giving simpler recovery and higher yields, for treating bacterial  
 PT infections  
 PS Disclosure; Page 27-28; 38pp; English.  
 CC Human bactericidal/permeability increasing protein (an endotoxin  
 CC binding protein) is useful for treatment of bacterial infections.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 37; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KELKRIKI 8  
 |||||  
 Db 58 KELKRIKI 65

RESULT 11  
 R57901  
 ID R57901 standard; Protein; 487 AA.  
 AC R57901;  
 DT 29-MAR-1995 (first entry)  
 DE Bactericide BPI.  
 KW BPI; bactericidal-permeability-increasing protein; bactericide;  
 KW fusion protein; Gram-negative bacterium; infection.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..31  
 FT /label= Sig\_peptide  
 PN W09418323-A.  
 PD 18-AUG-1994.  
 PF 02-FEB-1994; U01235.  
 PR 02-FEB-1993; US-013801.

CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxaemia and endotoxic shock.  
 SQ Sequence 456 AA;

Query Match 100.0%; Score 37; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
 |||||  
 Db 27 KELKRIKI 34

## RESULT 6

R24777 R24777 standard; Protein: 487 AA.  
 AC 02-JAN-1992 (first entry)  
 DE Sequence of human bactericidal/permeability-increasing protein  
 DE (hBPI) derived from human promyelocytic leukemia cells (HL-60).  
 KW Bactericidal; endotoxin inhibitor; anti-microbial.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..31  
 FT modified\_site 153..155  
 FT modified\_site /note= "potential N-linked glycosylation site"  
 FT modified\_site 379..381  
 FT /note= "see above"  
 PN W09209621-A.  
 PD 11-JUN-1992.  
 PF 03-DEC-1991; U09033.  
 PR 03-DEC-1990; US-621473.  
 PR 26-AUG-1991; US-754204.  
 PA (UYN) UNIV NEW YORK STATE.  
 PI Elsbach P, Weiss J;  
 PI WPI: 92-217016/26.  
 DR N-PSDB; Q25593.  
 PT Antibacterial fragments of bactericidal-permeability-increasing  
 PT protein - for treating Gram negative bacterial infections esp. in  
 PT immuno suppressed patients, with endotoxin activity  
 PS Claim 16; Fig 5; 63pp; English.  
 CC The hBPI holoprotein was isolated from human promyelocytic leukemia  
 CC cells (HL-60) and BPI cDNA was prep'd. The N-terminal fragments of  
 CC BPI exhibit all the anti-bacterial and membrane permeability-  
 CC increasing functions contained in the complete molecule (as well as  
 CC the LPS-neutralising function) but are substantially smaller. In  
 CC addition, the N-terminal fragment, as well as the C-terminal  
 CC fragment(s) are capable of inhibiting the prodn. and/or release of  
 CC tumour necrosis factor (TNF) triggered by LPS in whole blood. The  
 CC inventors claim the DNA sequence comprising bps 721-1491 of Q25593.  
 CC They also claim the following fragments of MATURE hBPI: 1-99 (I);  
 CC 1-203 (II); 200-456 (III); 204-456 (IV). NB: The above fragments  
 CC are labeled with respect to the first AA of the mature sequence,  
 CC which corresp. to residue 32 of R24777.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 37; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
 |||||  
 Db 58 KELKRIKI 65

## RESULT 7

R33776 R33776 standard; protein; 368 AA.

R33776;  
 22-JUL-1993 (first entry)  
 DT BPI from p337  
 DE Bacterial/permeability increasing protein; lipopolysaccharide; LPS;  
 KW inhibitor; endotoxin-related shock; disseminated intravascular  
 KW coagulation; anaemia; thrombocytopenia; ARDS; renal failure;  
 KW endotoxaemia; adult respiratory distress syndrome.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..32 "signal peptide"  
 FT region 32..368  
 FT /note= "mature protein"

PN W09305797-A.  
 PD 01-APR-1993.  
 PF 28-SEP-1992; U08234.  
 PR 27-SEP-1991; US-766566.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW, Snable JL, Wilde CG;  
 PI WPI: 93-117241/14.  
 DR Compsns. contg. a bactericidal and permeability increasing  
 PT protein - solubilised in a lipid carrier to maintain stably an  
 PT active form  
 PS Disclosure; Fig 6; 53pp; English.  
 CC The human bactericidal/permeability increasing protein may be used  
 CC in a compsn. solubilised in a lipid carrier. The BPI binds  
 CC lipopolysaccharides and inhibits the immunostimulatory and toxic  
 CC activities of LPS in vitro and in vivo. The lipid compsn. stably  
 CC maintains the BPI in a sol. active form. The compsn. can be used  
 CC for detecting or quantifying LPS in a sample, coating a surgical  
 CC tool, coating an implantable invasive device, decontaminating a  
 CC fluid contg. LPC and treating endotoxin related shock,  
 CC -disseminated intravascular coagulation, -anaemia, -thrombocytopenia,  
 CC -adult respiratory stress syndrome or -renal failure, or for  
 CC preventing endotoxaemia.  
 CC See also R33777, R33729-31.  
 SQ Sequence 368 AA;

Query Match 100.0%; Score 37; DB 1; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
 |||||  
 Db 58 KELKRIKI 65

## RESULT 8

R33777 R33777 standard; protein; 243 AA.

AC R33777;  
 DT 22-JUL-1993 (first entry)  
 DE BPI from p212.  
 KW Bacterial/permeability increasing protein; lipopolysaccharide; LPS;  
 KW inhibitor; endotoxin-related shock; disseminated intravascular  
 KW coagulation; anaemia; thrombocytopenia; ARDS; renal failure;  
 KW endotoxaemia; adult respiratory distress syndrome.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..32  
 FT region 32..243  
 FT /note= "signal peptide"  
 FT /note= "mature protein"

PN W09305797-A.  
 PD 01-APR-1993.  
 PF 28-SEP-1992; U08234.  
 PR 27-SEP-1991; US-766566.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW, Snable JL, Wilde CG;  
 PI WPI: 93-117241/14.  
 DR Compsns. contg. a bactericidal and permeability increasing  
 PT protein - solubilised in a lipid carrier to maintain stably an

FT region 1. 31  
 FT WO9203535-A.  
 PN 05-MAR-1992.  
 PD 13-AUG-1991: U05758.  
 PF 13-AUG-1990: US-567016.  
 PR 05-APR-1991: US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI: 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Disclosure: Fig 13; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxaemia and endotoxic shock.  
 SQ Sequence 368 AA;

Query Match 100.0%; Score 37; DB 1; Length 368;  
 Best Local Similarity 100.0%; Pred. NO. 5.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 DB 58 KELKRIKI 65

RESULT 3  
 R21841  
 ID R21841 standard; Protein: 243 AA.  
 AC R21841;  
 DT 22-JUL-1992 (first entry)  
 DE Sequence of p212, a biologically active fragment of the human  
 DE 57kd protein which binds to the outer membrane of susceptible gram  
 DE negative bacteria (BPI).  
 KW Endotoxin binding protein; diagnosis; therapy; endotoxaemia;  
 KW endotoxic shock.  
 OS Homo sapiens.  
 FH key  
 FT region 1. 31  
 FT Location/Qualifiers  
 FT 1. 31  
 FT /label= residues -31 to -1  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991: U05758.  
 PR 13-AUG-1990: US-567016.  
 PR 05-APR-1991: US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI: 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Disclosure: Fig 14; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxaemia and endotoxic shock.  
 SQ Sequence 243 AA;

Query Match 100.0%; Score 37; DB 1; Length 243;  
 Best Local Similarity 100.0%; Pred. NO. 3.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 DB 58 KELKRIKI 65

RESULT 4  
 R21843  
 ID R21843 standard; Protein: 396 AA.  
 AC R21843;  
 DT 22-JUL-1992 (first entry)  
 DE Sequence of CHO-BPI a biologically active variant of BPI, a human  
 DE 57kd protein which binds to the outer membrane of susceptible gram  
 DE negative bacteria (BPI).  
 KW Endotoxin binding protein; diagnosis; therapy; endotoxaemia;  
 KW endotoxic shock.  
 OS Homo sapiens.  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991: U05758.  
 PR 13-AUG-1990: US-567016.  
 PR 05-APR-1991: US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI: 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Claim 5; Fig 24; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxaemia and endotoxic shock.  
 SQ Sequence 396 AA;

Query Match 100.0%; Score 37; DB 1; Length 396;  
 Best Local Similarity 100.0%; Pred. NO. 5.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 DB 27 KELKRIKI 34

RESULT 5  
 R21844  
 ID R21844 standard; Protein: 456 AA.  
 AC R21844;  
 DT 22-JUL-1992 (first entry)  
 DE Sequence of BPI (DP linkage), a biologically active variant of BPI, a  
 DE 57kd protein which binds to the outer membrane of susceptible gram  
 DE negative bacteria (BPI).  
 KW Endotoxin binding protein; diagnosis; therapy; endotoxaemia;  
 KW endotoxic shock.  
 OS Homo sapiens.  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991: U05758.  
 PR 13-AUG-1990: US-567016.  
 PR 05-APR-1991: US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI: 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Claim 52; Fig 25; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the

Anti-fungal BPI peptide  
rBPI(1-199)-Ser135

Run on: September 7, 1999, 22:49:56 ; Search time 135.78 Seconds  
(without alignments)  
1.396 Million cell updates/sec

29-JUN-1990 (first entry)  
Human bactericidal/permeability-increasing (BPI) holoprotein  
Polypeptide fragment; bacteraemia; sepsis; gram-negative bacteria;  
bactericidal/permeability-increasing (BPI) holoprotein;  
mammalian myeloid blood cells; polymorphonuclear leukocytes (PMN);  
human promyelocytic leukaemia cells (HL-60).

Homo Sapiens.	Location/Qualifiers
Key	

peptide	1. .31
protein	32. .487
region	153. .155
/note="potential glycosylation site"	
region	379. .381
/note="potential glycosylation site"	

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	487	1	P94636 Human bactericidal
2	37	100.0	368	1	R21840 Sequence of p337,
3	37	100.0	243	1	R21841 Sequence of p212,
4	37	100.0	396	1	R21843 Sequence of CHO-Bp
5	37	100.0	456	1	R21844 Sequence of BPI (D
6	37	100.0	487	1	R24777 Sequence of human
7	37	100.0	368	1	R33776 BPI from p337. Co
8	37	100.0	243	1	R33777 BPI from p212. Co
9	37	100.0	487	1	R44420 pING4512 encoded f
10	37	100.0	487	1	R43886 Human bactericidal
11	37	100.0	487	1	R57901 Bactericide BPI. B
12	37	100.0	487	1	R23370 Recombinant bacter
13	37	100.0	15	1	R63686 BPI derived peptid
14	37	100.0	22	1	R63684 BPI derived peptid
15	37	100.0	29	1	R63682 BPI domain I. New
16	37	100.0	487	1	R62344 Recombinant 23 kD
17	37	100.0	487	1	R67998 Holoprotein rBP150
18	37	100.0	487	1	R71290 Human bactericidal
19	37	100.0	368	1	R71291 Human bactericidal
20	37	100.0	243	1	R71292 BPI(1-159)/lipopol.
21	37	100.0	199	1	R68925 BPI(1-85)/lipopoly
22	37	100.0	199	1	R68932 BPI(1-147)/LBP(147
23	37	100.0	198	1	R68933 BPI(1-85)/LBP(86-9
24	37	100.0	196	1	R68934 BPI(1-137)/lipopol
25	37	100.0	198	1	R68927 Bactericidal/perme
26	37	100.0	487	1	R68923 Recombinant bacter
27	37	100.0	487	1	R79185 Recombinant BPI ho
28	37	100.0	487	1	R76243 Bacterial permeabi
29	37	100.0	29	1	R76244 Bacterial permeabi
30	37	100.0	30	1	R76245 Bacterial permeabi
31	37	100.0	22	1	R76246 Bacterial permeabi
32	37	100.0	15	1	R76248 Bacterial permeabi
33	37	100.0	487	1	R79021 Bactericidal/perme
34	37	100.0	30	1	R86457 BPI.14 for use in
35	37	100.0	22	1	R86458 BPI.4 for use in t
36	37	100.0	15	1	R86460 BPI.54 for use in
37	37	100.0	487	1	R81245 Bactericidal/perme
38	37	100.0	487	1	R81247 Recombinant bacter
39	37	100.0	29	1	R84556 BPI Domain I for u
40	37	100.0	487	1	R84555 rBPI for use in tr
41	37	100.0	15	1	R81136 Anti-fungal BPI pe
42	37	100.0	29	1	R81085 Bactericidal/perme
43	37	100.0	22	1	R81091 Antifungal BPI pe

Query Match 100.0%; Score 37; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 8; Conservative 0; Mismatches 0; Indels

QY 1 KELRIKI 8  
DB 58 KELRIKI 65

RESULT	2
R21840	
ID	R21840 standard; protein; 368 AA.
AC	R21840;
DT	22-JUL-1992 (first entry)
DE	Sequence of p337, a biologically active fragment of the human
DE	57kd protein which binds to the outer membrane of susceptible gram
DE	negative bacteria (BFI).
DE	Endotoxin binding protein; diagnosis; therapy; endotoxemia;
KW	endotoxic shock.
OS	Homo sapiens.
Key	
Location/Qualifiers	

**THIS PAGE BLANK (USPTO)**

Db 487 FIKPGAVRI 495

RESULT 15

Q41058 PRELIMINARY; PRT: 922 AA.

ID Q41058;

AC Q41058;

DT 01-NOV-1996 (TREMREL. 01, Created)

DT 01-NOV-1996 (TREMREL. 01, Last sequence update)

DT 01-JAN-1999 (TREMREL. 09, Last annotation update)

DE STARCH BRANCHING ENZYME I PRECURSOR.

GN SBEI.

OS Pisum sativum (Garden pea).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC eukaryotes; Spermatophyta; Magnoliophyta; eudicotyledons;

OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;

OC Pisum.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 95201826.

RA BURTON R.A., BEWLEY J.D., SMITH A.M., BHATTACHARYA M.K., TATGE H.,

RA RING S., BULL V., HAMILTON W.D.O., MARTIN C.;

RT "Starch branching enzymes belonging to distinct enzyme families are

RT differentially expressed during pea embryo development.";

RL Plant J. 7:3-15(1995).

DR EMBL; X80009; CAA56319.1; -

DR MENDEL; 16398; PISA; Sbel; mnl6398.

DR PFAM; PF00128; alpha-amylase; 1.

KW Transit peptide.

FT TRANSIT 1 48 POTENTIAL.

FT CHAIN 49 922 POTENTIAL.

SQ SEQUENCE 922 AA; 105227 MW; OEDED7C1 CRC32;

Query Match 73.9%; Score 34; DB 10; Length 922;

Best Local Similarity 75.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKPGKVYK 9

Db 788 LKPGKVYKI 795

Search completed: September 7, 1999, 22:47:33

Job time: 7953 sec

DT 01-AUG-1998 (TREMREL. 07, Last annotation update)

DE FGFR SIGNALLING ADAPTOR SNT-2.

GN SNT-2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-PLACENTA;

RA XU H., LEE K., GOLDFARB M.;

RA Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF036718; AAB92555.1; -

SQ SEQUENCE 492 AA; 54461 MW; 3B1EDAFD CRC32;

Query Match 73.9%; Score 34; DB 4; Length 492;

Best Local Similarity 75.0%; Pred. No. 1e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGVKYK 8

Db 234 FLQPGQVK 241

RESULT 14

O16580 PRELIMINARY; PRT: 566 AA.

ID O16580;

AC O16580;

DT 01-JAN-1998 (TREMREL. 05, Created)

DT 01-JAN-1998 (TREMREL. 05, Last sequence update)

DT 01-NOV-1998 (TREMREL. 08, Last annotation update)

DE C33C12.3 PROTEIN.

GN C33C12.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;

OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE; 94150718.

RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,

RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,

RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,

RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,

RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,

RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,

RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,

RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,

RA THIRRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,

RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RT elegans.";

RL Nature 368:32-38(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA GATTUNG S., SCHEET P.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA WATERSTON R.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF016662; AAB66056.1; -

SQ SEQUENCE 566 AA; 63692 MW; 7CBA2F17 CRC32;

Query Match 73.9%; Score 34; DB 5; Length 566;

Best Local Similarity 55.6%; Pred. No. 1.2e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLKPGKVYK 9

Db 1:|||||:



OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Solanum.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. AMANDA;  
 RX MEDLINE: 98278379.  
 RA LARSSON C.T., KHOSHNOODI J., EK B., RASK L., LARSSON H.;  
 RT "Molecular cloning and characterization of starch-branching enzyme II  
 from potato.";  
 RL Plant Mol. Biol. 37:503-511(1998).  
 CC -1- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF  
 CC GLYCOGEN.  
 DR EMBL: AJ000004; CAA03846.1; -.  
 DR PFAM: PF00128; alpha-amylase; 1.  
 KW Transferase; Glycosyltransferase; Plasmid.  
 FT NON\_TER 1  
 SQ SEQUENCE 830 AA; 94967 MW; A3DD8A8 CRC32;

Query Match 76.1%; Score 35; DB 10; Length 830;  
 Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LKPGKVKV 9  
 ||||| ||  
 Db 752 LKPGKVKV 759

## RESULT 10

Q040663 ID Q040663 PRELIMINARY; PRT; 825 AA.  
 AC Q40663;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
 DE BRANCHING ENZYME-3 PRECURSOR.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Oryza.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 93366833.  
 RA MIZUNO K., KAWASAKI T., SHIMADA H., SATOH H., KOBAYASHI E.,  
 RA OKUMURA S., ARAI Y., BABA T.;  
 RT "Alteration of the structural properties of starch components by the  
 RT lack of an isoform of starch branching enzyme in rice seeds.";  
 RL J. Biol. Chem. 268:19084-19091(1993).  
 DR EMBL: D16201; BAA03738.1; -.  
 DR MENDEL; 16397; Oryza; Sbel; mn16397.  
 DR PFAM: PF00128; alpha-amylase; 1.  
 KW Transit peptide.  
 FT TRANSIT 1 65 POTENTIAL.  
 FT CHAIN 66 825 BRANCHING ENZYME-3.  
 SQ SEQUENCE 825 AA; 92757 MW; 00A9E6A0 CRC32;

Query Match 76.1%; Score 35; DB 10; Length 825;  
 Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LKPGKVKV 9  
 ||||| ||  
 Db 767 LKPGKVKV 774

## RESULT 11

Q92TB6 ID Q92TB6 PRELIMINARY; PRT; 829 AA.  
 AC Q92TB6;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE STARCH BRANCHING ENZYME IIB.  
 GN SBEIIB.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. BOMI;  
 RX MEDLINE: 98404232.  
 RA SUN C., SATHISH P., AHLANDSBERG S., JANSSON C.;  
 RT "The two genes encoding starch-branching enzymes Iia and Iib are  
 RT differentially expressed in barley.";  
 RL Plant Physiol. 118:37-49(1998).  
 DR EMBL: AF064561; AAC69754.1; -.  
 SQ SEQUENCE 829 AA; 93602 MW; 27277800 CRC32;

Query Match 76.1%; Score 35; DB 10; Length 829;  
 Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LKPGKVKV 9  
 ||||| ||  
 Db 771 LKPGKVKV 778

## RESULT 12

Q91132 ID Q91132 PRELIMINARY; PRT; 1642 AA.  
 AC Q91132;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE COBRA VENOM FACTOR PRECURSOR.  
 OS Naja naja (Indian cobra).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Lepidosauria;  
 OC Squamata; Scleroglossa; Serpentes; Colubroidea; Elapidae; Elapinae;  
 OC Naja.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=VENOM GLAND;  
 RX MEDLINE: 95108041.  
 RA FRITZINGER D.C., BREDEHORST R., VOGEL C.W.;  
 RT "Molecular cloning and derived primary structure of cobra venom  
 RT factor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12775-12779(1994).  
 DR EMBL: U09969; AAA68989.1; -.  
 DR PFAM: PF00207; A2M; 1.  
 DR PROSITE: PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 KW Signal.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 648 ALPHA CHAIN OF COBRA VENOM FACTOR.  
 FT CHAIN 1264 1642 BETA CHAIN OF COBRA VENOM FACTOR.  
 SQ SEQUENCE 1642 AA; 184516 MW; DC8042EA CRC32;

Query Match 76.1%; Score 35; DB 13; Length 1642;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
 ||: || ||||  
 Db 1450 FIQPGSVKV 1458

## RESULT 13

Q43559 ID Q43559 PRELIMINARY; PRT; 492 AA.  
 AC Q43559;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

Wed Sep 8 09:43:25 1999

RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,  
 RA OVERBEER R., COCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,  
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KATNE B.P., SYKES S.M.,  
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESSE C.R.,  
 RA VENTER J.C.;  
 RA "The complete genome sequence of the hyperthermophilic,  
 RT sulphate-reducing archaeon *Archaeoglobus fulgidus*.";  
 RL Nature 350:364-370(1997).  
 DR EMBL; AF001036; AAB90254.1; -  
 DR TIGR; AF0989; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 643 AA; 70844 MW; 3E7A0C3C CRC32;

Query Match 76.1%; Score 35; DB 1; Length 643;  
 Best Local Similarity 44.4%; Pred. No. 85;  
 Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLKPGKVKV 9  
 DB 320 YVKGRIKI 328  
 :|||||:

RESULT 8  
 O96270 PRELIMINARY; PRT; 338 AA.  
 ID O96270  
 AC O96270  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE PSEUDOURIDINE SYNTHETASE (RSUA FAM.), 1ST EUK. MEMBER (OO).  
 GN PFB0890C.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA KOONTN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
 RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
 RA "Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*."  
 RT Science 282:1126-1132(1998).  
 RL EMBL; AF001423; AAC71967.1; -  
 SQ SEQUENCE 338 AA; 40297 MW; AFD23C5 CRC32;

Query Match 76.1%; Score 35; DB 5; Length 338;  
 Best Local Similarity 66.7%; Pred. No. 44;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLKPGKVKV 9  
 DB 66 FIKDGKVKI 74  
 :|||||:

RESULT 9  
 O49953 PRELIMINARY; PRT; 830 AA.  
 ID O49953  
 AC O49953  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-WAY-1999 (Tremblrel. 10, Last annotation update)  
 DE STARCH BRANCHING ENZYME II, SBE-II (EC 2.4.1.18)  
 DE (1,4-ALPHA-GLUCAN BRANCHING ENZYME) (GLYCOGEN BRANCHING ENZYME)  
 DE (AMYLO-(1,4 TO 1,6)-TRANSGLUCOSIDASE)  
 DE (AMYLO-(1,4-1,6)-TRANSGLUCOSYLASE) (FRAGMENT).  
 GN SBE-II.  
 OS Solanum tuberosum (Potato).  
 OG Plasmid pALSBE-II (#123).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

RESULT 6  
 O16581 PRELIMINARY; PRT; 561 AA.  
 ID O16581  
 AC O16581  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
 DE C3C12.8 PROTEIN.  
 GN C3C12.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,  
 RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., JIER M., JOHNSTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., LAISTER N., LATREILLE P.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTON B., O'CALLAGHAN M.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*."  
 RT Nature 368:32-38(1994).  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA GAITUNG S., SCHEET P.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF016662; AAB66060.1; -  
 SQ SEQUENCE 561 AA; 62852 MW; AE8C083E CRC32;

Query Match 78.3%; Score 36; DB 5; Length 561;  
 Best Local Similarity 66.7%; Pred. No. 48;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLKPGKVKV 9  
 DB 487 FLKPGAVRI 495  
 :|||||:

RESULT 7  
 O29273 PRELIMINARY; PRT; 643 AA.  
 ID O29273  
 AC O29273  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-AUG-1998 (Tremblrel. 07, Last annotation update)  
 DE CONSERVED HYPOTHETICAL PROTEIN.  
 GN AF0989.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 OC Archaeoglobus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE; 98049343.  
 RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,  
 RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRIDES N.C.,  
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,

RX MEDLINE; 98241545.  
 RA SUGIYAMA T., KIDO N., KATO Y., KOIDE N., YOSHIDA T., YOKOCHI T.;  
 RT "Generation of *Escherichia coli* O9a serotype, a subtype of *E. coli*  
 RT O9, by transfer of the wb\* gene cluster of *Klebsiella O3* into *E. coli*  
 RT via recombination.";  
 RL J. Bacteriol. 180:2775-2778(1998).  
 DR EMBL; AB010293; BAA28334.1; -.  
 DR PFAM; PF00408; PGM\_PMM; 1.  
 DR PROSITE; PS00710; PGM\_PMM; 1.  
 KW Isomerase.  
 SQ SEQUENCE 460 AA; 50952 MW; FDBB7BAC CRC32;

Query Match 82.6%; Score 38; DB 2; Length 460;  
 Best Local Similarity 77.8%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
 |||||:|  
 Db 35 FLKPGKIVV 43

RESULT 3  
 O66240 PRELIMINARY; PRT; 460 AA.  
 ID O66240;  
 AC O66240;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE PHOSPHOMANNOMUTASE (EC 5.4.2.8).  
 GN MANB.  
 OS *Escherichia coli*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BI316-42 SEROTYPE O9;  
 RX MEDLINE; 98241545.  
 RA SUGIYAMA T., KIDO N., KATO Y., KOIDE N., YOSHIDA T., YOKOCHI T.;  
 RT "Generation of *Escherichia coli* O9a serotype, a subtype of *E. coli*  
 RT O9, by transfer of the wb\* gene cluster of *Klebsiella O3* into *E. coli*  
 RT via recombination.";  
 RL J. Bacteriol. 180:2775-2778(1998).  
 DR EMBL; AB010294; BAA28334.1; -.  
 DR PFAM; PF00408; PGM\_PMM; 1.  
 DR PROSITE; PS00710; PGM\_PMM; 1.  
 KW Isomerase.  
 SQ SEQUENCE 460 AA; 51174 MW; 059BD508 CRC32;

Query Match 82.6%; Score 38; DB 2; Length 460;  
 Best Local Similarity 77.8%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
 |||||:|  
 Db 35 FLKPGKIVV 43

RESULT 4  
 O66243 PRELIMINARY; PRT; 460 AA.  
 ID O66243;  
 AC O66243;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE PHOSPHOMANNOMUTASE (EC 5.4.2.8).  
 GN MANB.  
 OS *Klebsiella pneumoniae*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Klebsiella.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-K53 SEROTYPE O3;  
 RX MEDLINE; 98241545.  
 RA SUGIYAMA T., KIDO N., KATO Y., KOIDE N., YOSHIDA T., YOKOCHI T.;  
 RT "Generation of *Escherichia coli* O9a serotype, a subtype of *E. coli*  
 RT O9, by transfer of the wb\* gene cluster of *Klebsiella O3* into *E. coli*  
 RT via recombination.";  
 RL J. Bacteriol. 180:2775-2778(1998).  
 DR EMBL; AB010295; BAA28338.1; -.  
 DR PFAM; PF00408; PGM\_PMM; 1.  
 DR PROSITE; PS00710; PGM\_PMM; 1.  
 KW Isomerase.  
 SQ SEQUENCE 460 AA; 50712 MW; 1529D583 CRC32;

Query Match 82.6%; Score 38; DB 2; Length 460;  
 Best Local Similarity 77.8%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
 |||||:|  
 Db 35 FLKPGKIVV 43

RESULT 5  
 Q59428 PRELIMINARY; PRT; 460 AA.  
 ID Q59428;  
 AC Q59428;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE PHOSPHOMANNOMUTASE (EC 5.4.2.8).  
 GN RFBK OR MANB.  
 OS *Escherichia coli*, and *Klebsiella pneumoniae*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-F719;  
 RX MEDLINE; 94214678.  
 RA SUGIYAMA T., KIDO N., KOMATSU T., OHTA M., JANN K., JANN B., SAEKI A.,  
 RA KATO N.;  
 RT "Genetic analysis of *Escherichia coli* O9 rfb: identification and DNA  
 RT sequence of phosphomannomutase and GDP-mannose pyrophosphorylase  
 RT genes.";  
 RL Microbiology 140:59-71(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K49S SEROTYPE O3;  
 RX MEDLINE; 98241545.  
 RA SUGIYAMA T., KIDO N., KATO Y., KOIDE N., YOSHIDA T., YOKOCHI T.;  
 RT "Generation of *Escherichia coli* O9a serotype, a subtype of *E. coli*  
 RT O9, by transfer of the wb\* gene cluster of *Klebsiella O3* into *E. coli*  
 RT via recombination.";  
 RL J. Bacteriol. 180:2775-2778(1998).  
 CC -1- CATALYTIC ACTIVITY: D-MANNOSE 1-PHOSPHATE -> D-MANNOSE  
 CC 6-PHOSPHATE.  
 DR EMBL; D43637; BAA07746.1; -.  
 DR EMBL; AB010296; BAA28342.1; -.  
 DR PFAM; PF00408; PGM\_PMM; 1.  
 DR PROSITE; PS00710; PGM\_PMM; 1.  
 KW Isomerase.  
 SQ SEQUENCE 460 AA; 50811 MW; 00201807 CRC32;

Query Match 82.6%; Score 38; DB 2; Length 460;  
 Best Local Similarity 77.8%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
 |||||:|  
 Db 35 FLKPGKIVV 43

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:31 ; Search time 160.2 Seconds  
(without alignments)  
3.458 Million cell updates/sec

Title: US-09-124-280A-24

Perfect score: 46

Sequence: 1 FLKPGKVKV 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL\_10:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	82.6	458	2	O66229
2	38	82.6	460	2	O66236
3	38	82.6	460	2	O66240
4	38	82.6	460	2	O66243
5	38	82.6	460	2	O59428
6	36	78.3	561	5	O16581
7	35	76.1	643	1	O39273
8	35	76.1	338	5	O96270
9	35	76.1	830	10	O49953
10	35	76.1	825	10	O40663
11	35	76.1	829	10	O92TB6
12	35	76.1	1642	13	O91132
13	34	73.9	1492	4	O43559
14	34	73.9	566	5	O16580
15	34	73.9	922	10	O41058
16	34	73.9	247	11	O35342
17	34	73.9	247	11	O08732
18	34	73.9	141	12	O69421
19	34	73.9	1143	12	O12817
20	33	71.7	103	2	O66628
21	33	71.7	1019	2	O68770
22	33	71.7	214	4	O92882
23	33	71.7	1657	5	P90761
24	33	71.7	478	5	O96750
25	33	71.7	2343	6	O18806
26	33	71.7	515	10	O23614
27	33	71.7	564	10	O92TU9
28	33	71.7	500	10	O92P02
29	33	71.7	251	11	O62422

30 32 69.6 627 1 O28241  
31 32 69.6 645 2 O51539  
32 32 69.6 464 2 O87950  
33 32 69.6 261 2 O92EAL  
34 32 69.6 128 3 O74550  
35 32 69.6 336 3 O94150  
36 32 69.6 347 4 O16081  
37 32 69.6 264 4 O75937  
38 32 69.6 1696 5 O21436  
39 32 69.6 337 5 P91828  
40 32 69.6 264 10 O82638  
41 32 69.6 1465 10 Q92VY9  
42 32 69.6 528 10 Q92VY8  
43 32 69.6 238 11 Q921D3  
44 32 69.6 293 12 O89458  
45 31 67.4 130 11 P97591

O28241 archaeoglob  
O51539 borrelia bu  
O87950 serratia ma  
O92eal rickettsia  
O74550 schizosacch  
O94150 candida alb  
Q16081 homo sapien  
O75937 homo sapien  
Q21436 caenorhabdi  
P91828 caenorhabdi  
O82638 arabidopsis  
Q92vy9 arabidopsis  
Q92vy8 arabidopsis  
Q921d3 rattus norv  
O89458 heliothis a  
P97591 rattus norv

## ALIGNMENTS

## RESULT 1

O66229 ID O66229 PRELIMINARY; PRT; 458 AA.  
AC O66229;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE PHOSPHOMANNOMUTASE (EC 5.4.2.8).  
GN MANB.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-F492;  
RX MEDLINE: 98241545.  
RA SUGIYAMA T., KIDO N., KATO Y., KOIDE N., YOSHIDA T., YOKOCHI T.;  
RT "Generation of Escherichia coli O9a serotype, a subtype of E. coli  
RT O9, by transfer of the wb\* gene cluster of Klebsiella O3 into E. coli  
RT via recombination";  
RL J. Bacteriol. 180:2775-2778(1998).  
DR EMBL: AB010150; BAA28323.1; -  
DR PFAM: PF00408; PGM\_PMM; 1.  
DR PROSITE: PS00710; PGM\_PMM; 1.  
KW Isomerase.  
SQ SEQUENCE 458 AA; 50662 MW; A2D6F9ED CRC32;

Query Match 82.6%; Score 38; DB 2; Length 458;  
Best Local Similarity 77.8%; Pred. No. 15;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

## QY 1 FLKPGKVKV 9

|||||: 1  
35 FLKPGKIVV 43

## RESULT 2

O66236 ID O66236 PRELIMINARY; PRT; 460 AA.  
AC O66236;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE PHOSPHOMANNOMUTASE (EC 5.4.2.8).  
GN MANB.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-N24C SEROTYPE O9a;

RX MEDLINE; 93252283.  
RA HU W.S., WANG R.Y.-H., SHIH J.W.-K., LO S.-C.;  
RT "Identification of a putative kfc-rpm1-rpl1 operon flanked by long  
RL inverted repeats in Mycoplasma fermentans (incognitus strain).";  
RL GENE 127:79-85(1993).  
CC -!- FUNCTION: IF-3 BINDS TO THE 30S RIBOSOMAL SUBUNIT AND SHIFTS THE  
CC EQUILIBRIUM BETWEEN 70S RIBOSOMES AND THEIR 50S AND 30S SUBUNITS IN  
CC FAVOR OF THE FREE SUBUNITS, THUS ENHANCING THE AVAILABILITY OF 30S  
CC SUBUNITS ON WHICH PROTEIN SYNTHESIS INITIATION BEGINS.  
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: BELONGS TO THE IF-3 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M95046; G150149; -.  
DR PIR: JN0653; JN0653  
DR PROSITE; PS00938; IF3; 1.  
DR PFAM; PF00707; IF3; 1.  
DR HSP; P03000; ITIF.  
KW INITIATION FACTOR; PROTEIN BIOSYNTHESIS.  
SQ SEQUENCE 212 AA; 24486 MW; FF935272 CRC32;

Query Match 69.6%; Score 32; DB 1; Length 212;  
Best Local Similarity 55.6%; Pred. No. 38;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
Db 173 YLOPNKIKV 181

RESULT 15  
KIP2\_YEAST  
ID KIP2\_YEAST STANDARD; PRT; 706 AA.  
AC P28743;  
DT 01-DEC-1992 (REL. 24, CREATED)  
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE KINESIN-LIKE PROTEIN KIP2.  
GN KIP2 OR YPL155C OR P2581.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RX MEDLINE; 93317166.  
RA ROOF D.M., MELUH P.B., ROSE M.D.;  
RT "MRI and intraocular tamponade media.";  
RL J. CELL BIOL. 118:95-108(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RX MEDLINE; 97103777.  
RA PURNELLE B., COSTER F., GORFEAU A.;  
RT "The sequence of 55 kb on the left arm of yeast chromosome XVI  
RT identifies a small nuclear RNA, a new putative protein kinase and two  
RT new putative regulators.";  
RL YEAST 12:1483-1492(1996).  
CC -!- FUNCTION: REQUIRED FOR ASSEMBLY OF THE MITOTIC SPINDLE.  
CC -!- SUBUNIT: MIGHT BE DIMERIC.  
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC EMBL; Z11963; G3854; -.  
DR EMBL; Z73511; E247046; -.  
DR EMBL; X96770; E239045; -.  
DR PIR: C42640; C42640.  
DR SGD; L0000908; KIP2.  
DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
DR PROSITE; PS0067; KINESIN\_MOTOR\_DOMAIN2; 1.  
DR PFAM; PF00225; Kinesin; 2.  
DR HSP; P17119; 3KAR.  
KW MOTOR PROTEIN; CELL DIVISION; MICROTUBULES; ATP-BINDING; COILED COIL;  
KW MITOSIS; CELL CYCLE.  
FT DOMAIN 94 506 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).  
FT DOMAIN 507 541 COILED COIL (POTENTIAL).  
FT DOMAIN 569 589 COILED COIL (POTENTIAL).  
FT DOMAIN 612 689 COILED COIL (POTENTIAL).  
FT NP\_BIND 202 209 ATP (BY SIMILARITY).  
SQ SEQUENCE 706 AA; 78378 MW; A656E348 CRC32;

Query Match 69.6%; Score 32; DB 1; Length 706;  
Best Local Similarity 44.4%; Pred. No. 11e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
Db 75 FLHPGRIRI 83

Search completed: September 7, 1999, 23:59:05  
Job time: 517 sec

QY 1 FLKPGVKV 8  
| | | | |  
Db 140 FLKPGVKV 147

RESULT 12  
TUM\_BP186 STANDARD; PRT; 146 AA.  
ID TUM\_BP186  
AC P41063;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE SOS OPERON TUM PROTEIN (ORF95.1) [CONTAINS: ORF95.2; ORF95.4;  
ORF95.5].  
GN TUM OR CP95.  
OS BACTERIOPHAGE 186.  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; MYOVIRIDAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96204577.  
RA BRUMBY A.M., LAMONT I., DODD I.B., EGAN J.B.;  
RT "Defining the SOS operon of coliphage 186.";  
RL VIROLOGY 219:105-114(1996).  
RN [2]  
RP SEQUENCE OF 47-114 FROM N.A.  
RC STRAIN-186CITSP;  
RX MEDLINE: 90278955.  
RA SIVAPRASAD A.V., JARVINEN R., PUSPURA A., EGAN J.B.;  
RT "DNA replication studies with coliphage 186. III. A single phage gene  
is required for phage 186 replication.";  
RL J. MOL. BIOL. 213:449-463(1990).  
CC -!- FUNCTION: THE TUM PROTEIN IS RESPONSIBLE FOR UV INDUCTION OF THE  
186 PROPHAGE. THE OTHER THREE PROTEINS MODULATE TUM ACTIVITY.  
CC -!- ALTERNATIVE PRODUCTS: FOUR DIFFERENT PROTEINS (ORF95.1, ORF95.2  
ORF95.4 AND ORF95.5) ARE PRODUCED FROM THE SAME GENE BY THE USE  
OF ALTERNATIVE INITIATION SITES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U51471; G1262242;  
CC EMBL; U51471; G1262243;  
CC EMBL; U51471; G1262244;  
CC EMBL; U51471; G1262245;  
CC EMBL; X53318; G14902;  
CC PIR; S10633; S10633.  
CC HSP; P07737; IPFL.  
KW ALTERNATIVE INITIATION.  
FT CHAIN 1 146  
FT CHAIN 10 146  
FT CHAIN 47 146  
FT CHAIN 71 146  
FT CHAIN 12 12  
FT MUTAGEN 34 34  
FT MUTAGEN 40 40  
FT MUTAGEN 104 104  
FT SEQUENCE 146 AA; 16560 MW; 34C54D13 CRC32;  
Query Match 71.7%; Score 33; DB 1; Length 146;  
Best Local Similarity 77.8%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLKPGVKV 9

Db 65 FAKEGKVK 73  
| | | | |  
RESULT 13  
FRHD\_METJA STANDARD; PRT; 183 AA.  
ID FRHD\_METJA  
AC Q60339;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE COENZYME F420 HYDROGENASE DELTA SUBUNIT (PUTATIVE COENZYME F420  
DE HYDROGENASE PROCESSING SUBUNIT).  
GN FRHD OR MJ0030.  
OS METHANOCOCCUS JANNASCHII.  
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
OC METHANOCOCCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE: 96337999.  
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.L.,  
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLOCKER A.,  
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKY M.,  
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii.";  
RL SCIENCE 273:1058-1073(1996).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY U42. ALSO KNOWN AS THE  
HUPO/HYAD FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U67461; G1590832;  
CC TIGR; MJ0030;  
KW HYDROLASE; PROTEASE.  
SQ SEQUENCE 183 AA; 20628 MW; 5A9EB930 CRC32;  
Query Match 69.6%; Score 32; DB 1; Length 183;  
Best Local Similarity 50.0%; Pred. No. 33;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LKPGKVK 9  
: | | | | |  
Db 99 IKPGEIKI 106

RESULT 14  
IF3\_MYCFE STANDARD; PRT; 212 AA.  
ID IF3\_MYCFE  
AC Q05426;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE TRANSLATION INITIATION FACTOR IF-3.  
GN INFC.  
OS MYCOPLASMA FERMENTANS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;  
OC MYCOPLASMATACEAE; MYCOPLASMA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-INCOGNITUS;  
QY 1 FLKPGKVK 9

CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).  
 CC -----

DR EMBL: X75450; G438058; -;  
 DR EMBL: X84707; G683460; -;  
 DR MIM: 601340; -;  
 DR PFAM: PF00018; SH3; 1.  
 KW GROWTH FACTOR; SIGNAL.  
 FT SIGNAL 1 24  
 FT CHAIN 25 131  
 FT MELANOMA DERIVED GROWTH REGULATORY  
 FT PROTEIN.  
 SQ SEQUENCE 131 AA; 14509 MW; F91DC3FA CRC32;

Query Match 73.9%; Score 34; DB 1; Length 131;  
 Best Local Similarity 87.5%; Pred. No. 10;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LKPGKVKV 9  
 Db 114 LKPGKVDV 121

RESULT 10  
 POL\_FOAMY  
 ID POL\_FOAMY STANDARD; PRT; 886 AA.  
 AC P14350;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);  
 DE ENDONUCLEASE].  
 GN POL.

OS HUMAN SPUMARETROVIRUS (FOAMY VIRUS).  
 OC VIRUSES; RETROVIRIDAE; RETROVIRIDAE; SPUMAVIRUS.  
 RN [1]

RP SEQUENCE OF 1-545 FROM N.A.  
 RX MEDLINE: 88188241.  
 RA MAURER B., BANNERT H., DARAI G., FLUEGEL R.M.;  
 RT "Analysis of the primary structure of the long terminal repeat and  
 RT the gag and pol genes of the human spumaretrovirus.";  
 RL J. VIROL. 62:1590-1597(1988).  
 RN [2]

RP SEQUENCE OF 544-886 FROM N.A.  
 RX MEDLINE: 88004420.

RA FLUEGEL R.M., RETHWILM A., MAURER B., DARAI G.;  
 RT "Nucleotide sequence analysis of the env gene and its flanking  
 RT regions of the human spumaretrovirus reveals two novel genes.";  
 RL EMBO J. 6:2077-2084(1987).

CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
 CC -----

DR EMBL: X05591; G61760; -;  
 DR EMBL: M19427; G808677; -;  
 DR EMBL: M54978; G330972; -;  
 DR PIR: B28880; GNLJSP.  
 DR PFAM: PF00075; RnaseH; 1.  
 DR PFAM: PF00078; Ivt; 1.  
 DR PFAM: PF00552; Integrase; 1.  
 DR HSSP: P03355; IMWL.  
 DR KW TRANSFERASE; RNA-DIRECTED DNA POLYMERASE; HYDROLASE; NUCLEASE;  
 KW ENDONUCLEASE; POLYPROTEIN.  
 SQ SEQUENCE 886 AA; 100360 MW; CFEE0C18 CRC32;

Query Match 73.9%; Score 34; DB 1; Length 886;  
 Best Local Similarity 77.8%; Pred. No. 59;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 FLAPGKVKV 9  
 Db 585 FLEDGKVKV 593

RESULT 11  
 CRG2\_RANTE  
 ID CRG2\_RANTE STANDARD; PRT; 169 AA.  
 AC P02531;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE GAMMA CRYSTALLIN II (FRAGMENT)  
 OS RANA TEMPORARIA (EUROPEAN COMMON FROG).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
 OC NEOBATRACHIA; RANOIDEA; RANOIDAE; RANINAE; RANA.  
 RN [1]

RP SEQUENCE OF 1-168 FROM N.A.  
 RX MEDLINE: 84237544.  
 RA TOMAREV S.I., ZINOVIEVA R.D., CHALOVKA P., KRAYEV A.S.,  
 RA SKRYABIN K.G., GAUSE G.G. JR.;  
 RT "Multiple genes coding for the frog eye lens gamma-crystallins.";  
 RL GENE 27:301-308(1984).  
 RN [2]  
 RP SEQUENCE OF 1-168 FROM N.A.  
 RA TOMAREV S.I., ZINOVIEVA R.D., DOLGILEVICH S.M., KRAYEV A.S.,  
 RA SKRYABIN K.G., GAUSE G.G. JR.;  
 RT "Gamma crystallins of the eye lens of the grass frog are coded by a  
 RT family of multiple nonallelic genes.";  
 RL DOKL. BIOCHEM. 273:388-391(1984).  
 RN [3]

RN REVISION TO 169.

RA TOMAREV S.I.;

RL SUBMITTED (SEP-1985) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- FUNCTION: CRYSTALLINS ARE THE DOMINANT STRUCTURAL COMPONENTS  
 CC OF THE VERTEBRATE EYE LENS.

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- DOMAIN: HAS A TWO-DOMAIN BETA-STRUCTURE, FOLDED INTO FOUR  
 CC VERY SIMILAR GREEK KEY MOTIFS.

CC -1- THERE ARE AT LEAST FOUR GENES CODING FOR NONIDENTICAL GAMMA  
 CC CRYSTALLINS IN THIS MULTIGENE FAMILY.

CC -1- SIMILARITY: BELONGS TO THE BETA/GAMMA-CRYSTALLIN FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
 CC -----

DR EMBL: K02264; G213658; -;

DR EMBL: X00775; G64301; -;

DR PTR: A02937; CYEGG2

DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA; 2.

DR PFAM: PF00030; Crystall; 2.

DR HSSP: P02526; 1GCS.

DR EYE LENS PROTEIN; MULTIGENE FAMILY; DUPLICATION.

FT NON\_TER 1 1

FT DOMAIN <1 34 MOTIF 1.

FT DOMAIN 35 78 MOTIF 2.

FT DOMAIN 79 82 CONNECTING PEPTIDE.

FT DOMAIN 83 123 MOTIF 3.

FT DOMAIN 124 169 MOTIF 4.

SQ SEQUENCE 169 AA; 20160 MW; C3851DFC CRC32;

Query Match 71.7%; Score 33; DB 1; Length 169;  
 Best Local Similarity 75.0%; Pred. No. 20;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; A6001051; G2649835; -.  
CC DR TIGR; AF0768; -.  
CC DR PFAM; PF00009; GTP\_EFTU; 1.  
CC KW INITIATION FACTOR; PROTEIN BIOSYNTHESIS; GTP-BINDING.  
CC NP\_BIND 20 27 GTP (POTENTIAL).  
CC SQ SEQUENCE 595 AA; 66680 MW; 38FEF54C CRC32;

Query Match 73.9%; Score 34; DB 1; Length 595;  
Best Local Similarity 62.5%; Pred. No. 41;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LKPGKVKV 9  
:|||||:  
Db 466 IKGKIKL 473

RESULT 8  
MCTI\_RAT STANDARD; PRT; 260 AA.  
ID MCTI\_RAT  
AC P09650;  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MAST CELL PROTEINASE I PRECURSOR (EC 3.4.21.39) (RMCP-I) (CHYMASE).  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-MAST CELLS;  
RA LUTZELSCHWAB C.; PEULER G.; AVESKOGH M.; HELLMAN L.;  
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE OF 10-260 FROM N.A.  
RX MEDLINE; 94183231.  
RA ROULEAU A.; GARBARG M.; SCHWARTZ J.C.; RUAT M.;  
RT "Molecular cloning of rat mast cell protease 1 and development of  
RT specific probes for its gene transcript."  
RL BIOCHEM. BIOPHYS. RES. COMMON. 199:593-602(1994).  
RN [3]  
RP SEQUENCE OF 21-247.  
RX MEDLINE; 88107630.  
RA LE TRONG H.; PARMELEE D.C.; WALSH K.A.; NEURATH H.; WOODBURY R.G.;  
RT "Amino acid sequence of rat mast cell protease I (chymase).";  
RL BIOCHEMISTRY 26:6988-6994(1987).  
RN [4]  
RP SEQUENCE OF 21-73.  
RX MEDLINE; 79074806.  
RA WOODBURY R.G.; EVERITT M.; SANADA Y.; KATUNUMA N.; LAGUNOFF D.;  
RA NEURATH H.;  
RT "A major serine protease in rat skeletal muscle: evidence for its  
RT mast cell origin."  
RL PROC. NATL. ACAD. SCI. U.S.A. 75:5311-5313(1978).  
CC -1- CATALYTIC ACTIVITY: SPECIFICITY SIMILAR TO CHYMOTRYPSIN.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY. STRONGEST TO OTHER GRANZYMES AND TO MAST CELL  
CC PROTEASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U67915; G1698710; -.

DR EMBL; S69206; G546015; -.  
DR PIR; A28007; A28007.  
DR PIR; A05320; A05320.  
DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
DR PROSITE; PS00135; TRYPsin\_SER; 1.  
DR PFAM; PF00089; trypsin; 1.  
DR HSP; P00770; 3RP2.  
KW HYDROLASE; SERINE PROTEASE; ZYMOGEN; SIGNAL; MULTIGENE FAMILY.  
FT SIGNAL 1 18  
FT PROPEP 19 20  
FT CHAIN 21 260  
FT ACT\_SITE 65 65  
FT ACT\_SITE 109 109  
FT ACT\_SITE 202 202  
FT DISULFID 50 66  
FT DISULFID 143 208  
FT DISULFID 174 187  
FT CONFLICT 56 58  
FT CONFLICT 70 71  
SQ SEQUENCE 260 AA; 28617 MW; 0BB6F36 CRC32;

Query Match 73.9%; Score 34; DB 1; Length 260;  
Best Local Similarity 85.7%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLKPGKV 7  
:|||||:  
Db 136 FLKPGKM 142

RESULT 9  
MIA\_HUMAN STANDARD; PRT; 131 AA.  
ID MIA\_HUMAN  
AC Q16674;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MELANOMA DERIVED GROWTH REGULATORY PROTEIN PRECURSOR (MELANOMA  
DE INHIBITORY ACTIVITY).  
GN MIA.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE; 95007612.  
RA BLESCH A.; BOSSERHOFF A.-K.; APPEL R.; BEHL C.; HESSDOERFER B.;  
RA SCHMITT A.; JACHIMCZAK P.; LOTTSPETCH F.; BUETTNER R.; BOGDAN U.;  
RT "Cloning of a novel malignant melanoma-derived growth-regulatory  
RT protein, MIA.";  
RL CANCER RES. 54:5695-5701(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-PLACENTA;  
RX MEDLINE; 96132947.  
RA BOSSERHOFF A.-K.; HEIN R.; BOGDAN U.; BUETTNER R.;  
RT "Structure and promoter analysis of the gene encoding the human  
RT melanoma-inhibiting protein MIA.";  
J. BIOL. CHEM. 271:490-495(1996).  
CC -1- FUNCTION: ELICITS GROWTH INHIBITION ON MELANOMA CELLS IN VIRO  
CC AS WELL AS SOME OTHER NEUROECTODERMAL TUMORS, INCLUDING GLIOMAS.  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- TISSUE SPECIFICITY: ALL MALIGNANT MELANOMA CELL LINES TESTED AND  
CC INFREQUENTLY IN GLIOMA CELL LINES.

CC -1- PTM: MAY POSSESS TWO INTRAMOLECULAR DISULFIDE BONDS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



KW COMPLEMENT PATHWAY; COMPLEMENT ALTERNATE PATHWAY; PLASMA;  
 KW INFLAMMATORY RESPONSE; SIGNAL.

FT SIGNAL 1 22  
 FT CHAIN 23 1651 COMPLEMENT C3.  
 FT CHAIN 23 655 BETA CHAIN.  
 FT CHAIN 661 1651 ALPHA CHAIN.  
 FT PEPTIDE 661 738 C3A ANAPHYLATOXIN.  
 FT CHAIN 739 1651 C3B (ALPHA' CHAIN).  
 FT SITE 738 739 CLEAVAGE (BY C3 CONVERTASE).  
 FT DOMAIN 683 718 ANAPHYLATOXIN-LIKE.  
 FT DISULFID 545 807 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 615 650 BY SIMILARITY.  
 FT DISULFID 683 710 BY SIMILARITY.  
 FT DISULFID 684 717 BY SIMILARITY.  
 FT DISULFID 697 718 BY SIMILARITY.  
 FT DISULFID 863 1501 BY SIMILARITY.  
 FT DISULFID 1091 1147 BY SIMILARITY.  
 FT DISULFID 1346 1477 BY SIMILARITY.  
 FT DISULFID 1377 1446 BY SIMILARITY.  
 FT DISULFID 1494 1499 BY SIMILARITY.  
 FT DISULFID 1506 1578 BY SIMILARITY.  
 FT DISULFID 1525 1649 BY SIMILARITY.  
 FT THIOLEST 999 1002 BY SIMILARITY.  
 SQ SEQUENCE 1651 AA; 184926 MW; 2958575F CRC32;

Query Match 76.1%; Score 35; DB 1; Length 1651;  
 Best Local Similarity 66.7%; Pred. No. 68;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
 I::|||  
 DB 1459 FIQPGSVKV 1467

RESULT 6  
 RL27\_HUMAN  
 ID RL27\_HUMAN STANDARD; PRT; 135 AA.  
 AC P08526;  
 DT 01-AUG-1988 (REL. 08, CREATED)  
 DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE 60S RIBOSOMAL PROTEIN L27.  
 GN RPL27.  
 OS HOMO SAPIENS (HUMAN), RATTUS NORVEGICUS (RAT), AND  
 OS GALLUS GALLUS (CHICKEN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-HUMAN; TISSUE-KIDNEY;  
 RX MEDLINE; 94198298.  
 RA GALLAGHER R.A., MCCLEAN P.M., MALIK A.N.;  
 RT "Cloning and nucleotide sequence of a full length cDNA encoding  
 RT ribosomal protein L27 from human fetal kidney.";  
 RL BIOCHEM. BIOPHYS. ACTA 1217:329-332(1994).  
 RN [2]  
 RP SEQUENCE OF 1-97 FROM N.A.  
 RC SPECIES-HUMAN;  
 RA BHAT K.S.;  
 RL SUBMITTED (XXX-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-RAT; STRAIN-WISTAR; TISSUE-LIVER;  
 RX MEDLINE; 88185328.  
 RA TANAKA T., KUNANO Y., ISHIKAWA K., OGATA K.;  
 RT "Nucleotide sequence of cloned cDNA specific for rat ribosomal  
 RT protein L27.";  
 RL EUR. J. BIOCHEM. 173:53-56(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-CHICKEN;  
 RX MEDLINE; 91232917.

RA LEBEAU M.C., ALVAREZ-BOLADO G., BRAISSANT O., WAHLI W., CATSICAS S.;  
 RT "Ribosomal protein L27 is identical in chick and rat.";  
 RL NUCLEIC ACIDS RES. 19:1337-1337(1991).  
 CC -!- SIMILARITY: BELONGS TO THE L27E FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; L05094; G553642; -;  
 DR EMBL; L19527; G388769; -;  
 DR EMBL; X07424; G57694; -;  
 DR EMBL; X56852; G62981; -;  
 DR PIR; S00401; RSRT27.  
 DR PIR; S22288; S22288.  
 DR PIR; S43505; S43505.  
 DR PROSITE; PS01107; RIBOSOMAL\_L27E; 1.  
 KW RIBOSOMAL PROTEIN.  
 FT INIT\_MET 0  
 SQ SEQUENCE 135 AA; 15666 MW; 45DED05D CRC32;

Query Match 76.1%; Score 35; DB 1; Length 135;  
 Best Local Similarity 85.7%; Pred. No. 6.8;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKV 7  
 I:||||  
 DB 3 FMKPGKV 9

RESULT 7  
 IF2P\_ARCFU  
 ID IF2P\_ARCFU STANDARD; PRT; 595 AA.  
 AC Q29490;  
 DT 15-DEC-1998 (REL. 37, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE PROBABLE TRANSLATION INITIATION FACTOR IF-2.  
 GN INF2 OR AF0768.  
 OS ARCHAEOGLOBUS FULGIDUS.  
 OC ARCHAEA; EURYARCHAEOTA; ARCHAEOGLOBALES; ARCHAEOGLOBACEAE;  
 OC ARCHAEOGLOBUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE; 98049343.  
 RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,  
 RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRIDES N.C.,  
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
 RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,  
 RA OVERBECK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,  
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,  
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,  
 RA VENTER J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon Archaeoglobus fulgidus.";  
 RL NATURE 390:364-370(1997).  
 CC -!- FUNCTION: FUNCTION IN GENERAL TRANSLATION INITIATION BY PROMOTING  
 CC THE BINDING OF THE FORMYL METHIONINE-TRNA TO RIBOSOMES. SEEMS TO  
 CC FUNCTION ALONG WITH EIF-2 (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE IF-2 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC  
CC EMBL: L27646; G441137; -  
DR EMBL: L27632; G441140; -  
DR PROSITE: PS00710; PGM\_PMM; 1.  
DR PFAM: PF00408; PGM\_PMM; 1.  
KW LIPOPOLYSACCHARIDE BIOSYNTHESIS; ISOMERASE; PHOSPHORYLATION;  
MULTIGENE FAMILY.  
FT ACT\_SITE 98 FORMS THE PHOSPHOSERINE INTERMEDIATE  
(BY SIMILARITY).  
FT VARIANT 392 392 A -> R.  
FT VARIANT 441 441 T -> V.  
FT VARIANT 446 446 A -> E.  
FT VARIANT 449 453 EEILA -> TELLN.  
FT VARIANT 456 456 K -> KEELL.  
SQ SEQUENCE 456 AA; 50423 MW; 3AB76347 CRC32;

Query Match 82.6%; Score 38; DB 1; Length 456;  
Best Local Similarity 77.8%; Pred. No. 5.7; Mismatches 1; Indels 0; Gaps 0;  
Matches 7; Conservative 1;

QY 1 FLKPGKVKV 9  
| | | | | | | |  
Db 35 FLKPGKIVV 43

RESULT 4  
LBP\_RABIT STANDARD; PRT; 482 AA.  
ID LBP\_RABIT  
AC P17454;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).  
GN LBP.  
OS ORYCTOLAGUS CUNICULUS (RABBIT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
CC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 90385281.  
RA SCHUMANN R.R., LEONG S.R., FLAGGS G.W., GRAY P.W., WRIGHT S.D.,  
RA MATHISON J.C., TOBIAS P.S., ULEVITCH R.J.;  
RT "Structure and function of lipopolysaccharide binding protein.";  
RL SCIENCE 249:1429-1431(1990).  
RN [2]  
RP SEQUENCE OF 27-66.  
RC TISSUE-SERUM;  
RX TOBIAS P.S., SOLDAN K., ULEVITCH R.J.;  
RT "Isolation of a lipopolysaccharide-binding acute phase reactant from  
rabbit serum.";  
RL J. EXP. MED. 164:777-793(1986).  
CC -!- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL  
MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS  
TO INTERACT WITH THE CD14 RECEPTOR.  
CC -!- SIMILARITY: BELONGS TO THE BPL/CTP/LBP/PLTP FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: M35534; G165468; -  
CC PIR: B35843; B35843.  
DR PROSITE: PS00400; LBP\_BPL\_CETP; 1.  
DR PFAM: PF01273; LBP\_BPL\_CETP; 1.  
DR HSSP: P17213; LBPI.

LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.  
KW SIGNAL 1 26  
FT CHAIN 27 482 LIPOPOLYSACCHARIDE-BINDING PROTEIN.  
FT CARBOHYD 301 301 POTENTIAL.  
FT CARBOHYD 351 351 POTENTIAL.  
FT CARBOHYD 387 387 POTENTIAL.  
FT CONFLICT 57 57 E -> G (IN REF. 2).  
FT CONFLICT 63 63 S -> F (IN REF. 2).  
SQ SEQUENCE 482 AA; 54001 MW; 9DAFDF34 CRC32;

Query Match 78.3%; Score 36; DB 1; Length 482;  
Best Local Similarity 66.7%; Pred. No. 14; Mismatches 3; Indels 0; Gaps 0;  
Matches 6; Conservative 3;

QY 1 FLKPGKVKV 9  
| | | | | | | |  
Db 402 FLKPGRLQV 410

RESULT 5  
CO3\_NAJNA STANDARD; PRT; 1651 AA.  
ID CO3\_NAJNA  
AC Q01833;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE COMPLEMENT C3 PRECURSOR [CONTAINS: C3A ANAPHYLATOXIN].  
GN C3.  
OS NAJA NAJA (INDIAN COBRA).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;  
CC SCLEROGLOSSA; SERPENTES; COLUBROIDEA; ELAPIDAE; ELAPINAE; NAJA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE-LIVER;  
RX MEDLINE: 93056528.  
RA FRITZINGER D.C., CONNELLY M., PETRELLA E.C., BREDEHORST R.,  
RA VOGEL C.W.;  
RT "Primary structure of cobra complement component C3.";  
RL J. IMMUNOL. 149:3554-3562(1992).  
CC -!- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL  
REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.  
AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE  
THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.  
CC -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3, IT  
INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR  
PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND  
BASOPHILIC LEUKOCYTES.  
CC -!- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG  
RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE  
BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,  
CC RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA  
CHAIN).  
CC -!- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
CC -!- SIMILARITY: CONTAINS 1 ANAPHYLATOXIN-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: L02365; G213373; -  
CC PIR: A46513; A46513.  
DR PROSITE: PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
DR PROSITE: PS01177; ANAPHYLATOXIN\_1; 1.  
DR PROSITE: PS01178; ANAPHYLATOXIN\_2; 1.  
DR PFAM: PF00207; A2M; 1.  
DR HSSP: P01024; 1C3D.

```

DR EMBL: X98665; E264344; JOINED.
DR EMBL: X98666; E264344; JOINED.
DR EMBL: X98667; E264344; JOINED.
DR EMBL: X98668; E264344; JOINED.
DR EMBL: AF013512; G2653817; JOINED.
DR EMBL: AF013500; G2653817; JOINED.
DR EMBL: AF013501; G2653817; JOINED.
DR EMBL: AF013502; G2653817; JOINED.
DR EMBL: AF013503; G2653817; JOINED.
DR EMBL: AF013504; G2653817; JOINED.
DR EMBL: AF013505; G2653817; JOINED.
DR EMBL: AF013506; G2653817; JOINED.
DR EMBL: AF013507; G2653817; JOINED.
DR EMBL: AF013508; G2653817; JOINED.
DR EMBL: AF013509; G2653817; JOINED.
DR EMBL: AF013510; G2653817; JOINED.
DR EMBL: AF013511; G2653817; JOINED.
DR EMBL: L42172; G807652; -.
DR PIR: A35843; A35843.
DR MIM: 151990; -.
DR PROSITE: PS00400; LBP_BPI_CETP; 1.
DR PFAM: PF01273; LBP_BPI_CETP; 1.
DR HSSP: P17213; LBP1.
KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 481 LIPOPOLYSACCHARIDE-BINDING PROTEIN.
FT CARBOHYD 300 300 POTENTIAL.
FT CARBOHYD 355 355 POTENTIAL.
FT CARBOHYD 386 386 POTENTIAL.
FT CARBOHYD 394 394 POTENTIAL.
FT CONFLICT 82 82 N -> K (IN REF. 3).
FT CONFLICT 128 128 S -> F (IN REF. 3).
FT CONFLICT 154 157 VTAS -> GYCL (IN REF. 1).
FT CONFLICT 174 174 L -> S (IN REF. 1).
FT CONFLICT 257 257 R -> S (IN REF. 3).
FT CONFLICT 266 270 VMSLP -> A (IN REF. 1).
FT CONFLICT 369 369 L -> H (IN REF. 3).
FT CONFLICT 436 436 L -> F (IN REF. 3).
FT SEQUENCE 481 AA; 53349 MW; 9F3D0008 CRC32;

Query Match 100.0%; Score 46; DB 1; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
DB 401 FLKPGKVKV 409

RESULT 2
MANB_KLEPN STANDARD; PRT; 157 AA.
AC Q48463;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PHOSPHOMANNOMUTASE (EC 5.4.2.8) (PMK) (ORF17) (FRAGMENT).
GN MANB OR CPSS.
OS KLEBSIELLA PNEUMONIAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC KLEBSIELLA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95204345.
RA ARAKAWA Y., WACHAROTAYANKUN R., NAGATSUKA T., ITO H., KATO N.,
RA OHTA M.;
RT "Genomic organization of the Klebsiella pneumoniae cps region
RT responsible for serotype K2 capsular polysaccharide synthesis in the
RT virulent strain Chedid.";
RL J. BACTERIOL. 177:1788-1796(1995).
CC -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF THE K2 CAPSULAR

```

```

CC POLYSACCHARIDE BIOSYNTHESIS.
CC -!- CATALYTIC ACTIVITY: D-MANNOSE 1-PHOSPHATE -> D-MANNOSE 6-PHOSPHATE.
CC -!- PATHWAY: GDP-MANNOSE BIOSYNTHESIS (PART OF THE GENE CLUSTER FOR
CC THE K2 CAPSULAR POLYSACCHARIDE)
CC -!- SIMILARITY: TO OTHER PHOSPHOGLUCOMUTASES AND PHOSPHOMANNOMUTASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D21242; G74767; -.
CC PROSITE: PS00710; PGM_PMK; 1.
CC PFAM: PF00408; PGM_PMK; 1.
KW LIPOPOLYSACCHARIDE BIOSYNTHESIS; ISOMERASE; PHOSPHORYLATION.
FT ACT_SITE 98 FORMS THE PHOSPHOSERINE INTERMEDIATE
FT NON_TER 157 157 (BY SIMILARITY).
FT SEQUENCE 157 AA; 17234 MW; 1EA65F99 CRC32;

Query Match 82.6%; Score 38; DB 1; Length 157;
Best Local Similarity 77.8%; Pred. No. 2.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
DB 35 FLKPGKIV 43

RESULT 3
RFK9_ECOLI STANDARD; PRT; 456 AA.
AC P37755;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE PHOSPHOMANNOMUTASE (EC 5.4.2.8) (PMK).
GN RFBK.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-O9:K30:H12 / ISOLATE E69;
RX MEDLINE; 94252978.
RA JAYARATNE P., BRONNER D., MACLACHLAN R.P., DODGSON C., KIDO N.,
RA WHITFIELD C.;
RT "Cloning and analysis of duplicated rfbM and rfbK genes involved in
RT the formation of GDP-mannose in Escherichia coli O9:K30 and
RT participation of rfb genes in the synthesis of the group I K30
RT capsular polysaccharide.";
RL J. BACTERIOL. 176:3126-3139(1994).
CC -!- FUNCTION: INVOLVED IN GDP-MANNOSE BIOSYNTHESIS WHICH SERVES AS
CC THE ACTIVATED SUGAR NUCLEOTIDE PRECURSOR FOR MANNOSE RESIDUES
CC IN CELL SURFACE POLYSACCHARIDES. THIS ENZYME PARTICIPATES IN
CC SYNTHESIS OF THE LPS O9 ANTIGEN.
CC -!- CATALYTIC ACTIVITY: D-MANNOSE 1-PHOSPHATE -> D-MANNOSE 6-PHOSPHATE.
CC -!- PATHWAY: GDP-MANNOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
CC BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -!- THERE ARE TWO DUPLICATED GENES FOR RFBM AND RFBK IN THIS E.COLI
CC STRAIN.
CC -!- SIMILARITY: TO OTHER PHOSPHOGLUCOMUTASES AND PHOSPHOMANNOMUTASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:03 ; Search time 71.09 Seconds  
(without alignments)  
3.579 Million cell updates/sec

Title: US-09-124-280A-24  
Perfect score: 46  
Sequence: 1 FLKPKVKV 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	481	1	LBP_HUMAN
2	38	82.6	157	1	MANB_KLEPN
3	38	82.6	456	1	RFK9_ECOLI
4	36	78.3	482	1	LBP_RABIT
5	35	76.1	1651	1	CO3_NAJNA
6	35	76.1	135	1	RL27_HUMAN
7	34	73.9	595	1	IF2P_ARCFU
8	34	73.9	260	1	MCT1_RAT
9	34	73.9	131	1	MIA_HUMAN
10	34	73.9	886	1	POL_FOAMV
11	33	71.7	169	1	CRS2_RANTE
12	33	71.7	146	1	TUM_BPI86
13	32	69.6	183	1	FRHD_METJA
14	32	69.6	212	1	IF3_MYCFE
15	32	69.6	706	1	KIP2_YEAST
16	32	69.6	244	1	MCT2_MOUSE
17	32	69.6	246	1	MCT4_MOUSE
18	32	69.6	334	1	PSTS_HAEIN
19	32	69.6	136	1	RL27_YEAST
20	32	69.6	134	1	SMPB_TREPA
21	32	69.6	1081	1	SPS2_CRAPL
22	32	69.6	970	1	T3RE_BPI1
23	32	69.6	177	1	YB00_YEAST
24	31	67.4	510	1	BAG1_MOUSE
25	31	67.4	227	1	COX2_MACFA
26	31	67.4	758	1	CSTA_MYCTU
27	31	67.4	440	1	DNA4_THEMEA
28	31	67.4	1492	1	DPOA_OXYNO
29	31	67.4	588	1	GGT_MOUSE
30	31	67.4	568	1	GGT_RAT
31	31	67.4	384	1	HEP1_FLAHE
32	31	67.4	91	1	MIA_RAT
33	31	67.4	917	1	NIA2_ARATH
34	31	67.4	1054	1	POL_SIVMK
35	31	67.4	471	1	PTWB_BACST
36	31	67.4	136	1	R27A_SCHPO
37	31	67.4	136	1	R27B_SCHPO
38	31	67.4	134	1	RL27_PYRST
39	31	67.4	709	1	RRP2_INCUJ
40	31	67.4	1335	1	RRPO_FXWV
41	31	67.4	129	1	RS9_ECOLI
42	31	67.4	129	1	RS9_HAEIN
43	31	67.4	129	1	RS9_HAESO

44	31	67.4	28	1	RS9_THETH	P80374 thermus aqu
45	31	67.4	1121	1	YJ09_YEAST	P47107 saccharomyc
ALIGNMENTS						
RESULT 1						
LBP_HUMAN						
ID	LBP_HUMAN	STANDARD;	PRT;	481	AA.	
AC	P18428;	O92672;	O43438;			
DT	01-NOV-1990	(REL. 15,	CREATED)			
DT	15-DEC-1998	(REL. 37,	LAST SEQUENCE UPDATE)			
DT	15-DEC-1998	(REL. 37,	LAST ANNOTATION UPDATE)			
DE	LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).					
GN	LBP.					
OS	HOMO SAPIENS (HUMAN).					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;					
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE: 90385281.					
RA	SCHUMANN R.R., LEONG S.R., FLAGGS G.W., GRAY P.W., WRIGHT S.D.,					
RA	MATHISON J.C., TOBIAS P.S., ULEVITCH R.J.;					
RT	"Structure and function of lipopolysaccharide binding protein.";					
RL	SCIENCE 249:1429-1431(1990).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RA	HUBACEK J.A., ASLANIDIS C., SCHMITZ G.;					
RL	SUBMITTED (OCT-1996) TO EMBL/GENBANK/DDJB DATA BANKS.					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE: 98110577.					
RA	KIRSCHING C.J., AU-YOUNG J., LAMPING N., REUTER D., PFEIL D.,					
RA	SEILHAMER J.J., SCHUMANN R.R.;					
RT	"Similar organization of the lipopolysaccharide-binding protein (LBP)					
RT	and phospholipid transfer protein (PLTP) genes suggests a common gene					
RL	family of lipid-binding proteins.";					
RL	GENOMICS 46:416-425(1997).					
RN	[4]					
RP	SEQUENCE OF 1-41 FROM N.A.					
RA	SUTTON C.L., SMITH R.I.F., CENTOLA M.B., THEOFAN G.;					
RL	SUBMITTED (MAY-1995) TO EMBL/GENBANK/DDJB DATA BANKS.					
RN	[5]					
RP	3D-STRUCTURE MODELLING.					
RX	MEDLINE: 98227852.					
RA	BEAVER L.J., CARROLL S.F., EISENBERG D.;					
RT	"The BPI/LBP family of proteins: a structural analysis of conserved					
RL	regions.";					
RL	PROTEIN SCI. 7:906-914(1998).					
CC	-!- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL					
CC	LIPOPOLYSACCHARIDES (LPS). A GLYCOLIPID PRESENT IN THE OUTER					
CC	MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS					
CC	TO INTERACT WITH THE CD14 RECEPTOR.					
CC	-!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL; M35533; G186966;					
DR	EMBL; X98657; E264344;					
DR	EMBL; X98658; E264344; JOINED.					
DR	EMBL; X98659; E264344; JOINED.					
DR	EMBL; X98660; E264344; JOINED.					
DR	EMBL; X98661; E264344; JOINED.					
DR	EMBL; X98662; E264344; JOINED.					
DR	EMBL; X98663; E264344; JOINED.					
DR	EMBL; X98664; E264344; JOINED.					

**THIS PAGE BLANK (USPTO)**

Wed Sep 8 09:43:23 1999

Query Match 73.9%; Score 34; DB 2; Length 251;  
Best Local Similarity 85.7%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ELKPGKV 7  
Db 127 ELKPGNM 133

Search completed: September 7, 1999, 23:22:38  
Job time: 945 sec

Db 4 FMKPGKV 10

# RESULT 12

E69373

conserved hypothetical protein AF0989 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998

C:Accession: E69373

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343

A:Accession: E69373

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-643 <KLE>

A:Cross-references: GB:AE001036; GB:AE000782; NID:g2689359; PID:g2649609; TIGR:AF0989

Query Match 76.1%; Score 35; DB 2; Length 643;

Best Local Similarity 44.4%; Pred. No. 47;

Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLKPGKVKV 9

|||||:

Db 320 YVKPGRIKI 328

# RESULT 13

C71603

pseudouridine synthetase (RsuA fam.) PFB0890c - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998

C:Accession: C71603

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;

Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.

Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600

A:Accession: C71603

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-338 <GAR>

A:Cross-references: GB:AE001423; GB:AE001362; NID:g3845302; PID:g3845303; TIGR:PFB0890c

A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0890c

# Query Match

Best Local Similarity 76.1%; Score 35; DB 2; Length 338;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLKPGKVKV 9

|||||:

Db 66 FIKDGKVKI 74

# RESULT 14

GNLJSP

pol polyprotein - human foamy virus

C:Species: human foamy virus

C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 21-Nov-1997

C:Accession: A93033; A91074; A29685; B28880

R:Maurer, B.; Bannert, H.; Darai, G.; Fluegel, R.M.

J. Virol. 62, 1590-1597, 1988

A:Title: Analysis of the primary structure of the long terminal repeat and the gag and p

A:Reference number: A93033; MUID:88188241

A:Accession: A93033

A:Molecule type: genomic RNA

A:Residues: 1-545 <RSA>

A:Cross-references: GB:M19427; NID:g330974; PID:g808677

R:Fluegel, R.M.; Rethwilm, A.; Maurer, B.; Darai, G.

EMBO J. 6, 2077-2084, 1987

A:Title: Nucleotide sequence analysis of the env gene and its flanking regions of the

A:Reference number: A91074; MUID:88004420

A:Accession: A91074

A:Molecule type: genomic RNA

A:Residues: 544-886 <RSB>

C:Genetics:

A:Gene: pol

C:Superfamily: pol polyprotein

C:Keywords: polyprotein; reverse transcriptase

Query Match 73.9%; Score 34; DB 1; Length 886;

Best Local Similarity 77.8%; Pred. No. 98;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLKPGKVKV 9

|||||:

Db 585 FLEDGKVKV 593

# RESULT 15

JC2125

chymase (EC 3.4.21.39) precursor - rat

N:Alternate names: chymotrypsin-like proteinase; mast cell proteinase I; skeletal mus

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 29-Jan-1999

C:Accession: JC2125; A28007; A05320; P01825

R:Rouleau, A.; Garbarg, M.; Schwartz, J.C.; Ruat, M.

Biochem. Biophys. Res. Commun. 199, 593-602, 1994

A:Title: Molecular cloning of rat mast cell protease 1 and development of specific pr

A:Reference number: JC2125; MUID:94183231

A:Accession: JC2125

A:Molecule type: mRNA

A:Residues: 1-251 <ROU>

A:Cross-references: GB:S69206; NID:g546014; PID:g546015

R:Le Trong, H.; Parmelee, D.C.; Walsh, K.A.; Neurath, H.; Woodbury, R.G.

Biochemistry 26, 6988-6994, 1987

A:Title: Amino acid sequence of rat mast cell protease I (chymase).

A:Reference number: A28007; MUID:88107630

A:Accession: A28007

A:Molecule type: protein

A:Residues: 12-238 <LEP>

R:Woodbury, R.G.; Everitt, M.; Sanada, Y.; Katunuma, N.; Lagunoff, D.; Neurath, H.

Proc. Natl. Acad. Sci. U.S.A. 75, 5311-5313, 1978

A:Reference number: A05320; MUID:79074806

A:Accession: A05320

A:Molecule type: protein

A:Residues: 12-46, N', 48, N', 50-60, NE' <WOO>

A:Experimental source: peritoneal mast cells and skeletal muscle, residues 1-35, and

R:Nelson, R.B.; Siman, R.; Iqbal, M.A.; Potter, H.

J. Neurochem. 61, 567-577, 1993

A:Title: Identification of a chymotrypsin-like mast cell protease in rat brain capabl

A:Reference number: P01825; MUID:93329423

A:Accession: P01825

A:Molecule type: protein

A:Residues: 12-25, 27, 29, 31, 86-92, 223-229 <NEL>

A:Note: three peptide fragments match regions found in rat mast cell protease I

C:Comment: This protein is a specific marker of connective tissue mast cells selectiv

C:Comment: This enzyme has chymotrypsin-like specificity towards small substrates. It

C:Superfamily: trypsin; trypsin homology

C:Keywords: brain; glycoprotein; hydrolase; mast cell; serine proteinase; skeletal mu

F:1-11/Domain: signal sequence #status predicted <SIG>

F:12-251/Product: chymase #status predicted <WAT>

F:12-230/Domain: trypsin homology <TRY>

F:56,100,193/Active site: His, Asp, Ser #status predicted

R.Fritzinger, D.C.: Bredehorst, R.; Vogel, C.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 12775-12779, 1994  
A:Title: Molecular cloning and derived primary structure of cobra venom factor.  
A:Reference number: 151018; MUID:95108041  
A:Accession: 151018  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1642 <FRI>  
A:Cross-references: EMBL:U09969; NID:g881914; PID:g881915  
C:Superfamily: alpha-2-macroglobulin

Db 1459 FIOPGSVKV 1467

RESULT 7

R5RT27  
ribosomal protein L27, cytosolic - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 07-Nov-1997  
A:Accession: S00401; S11419  
R:Tanaka, T.; Kuwano, Y.; Ishikawa, K.; Ogata, K.  
Eur. J. Biochem. 173, 53-56, 1988  
A:Title: Nucleotide sequence of cloned cDNA specific for rat ribosomal protein L27.  
A:Reference number: S00401; MUID:88185328  
A:Accession: S00401  
A:Molecule type: mRNA  
A:Residues: 1-136 <FAN>  
A:Cross-references: EMBL:X07424; NID:g57693; PID:g57694  
R:Wittmann-Liebold, B.; Geissler, A.W.; Lin, A.; Wool, I.G.  
J. Supramol. Struct. 12, 425-433, 1979  
A:Title: Sequence of the amino-terminal region of rat liver ribosomal proteins S4, S6, S8  
A:Reference number: S11413; MUID:80252792  
A:Accession: S11419  
A:Molecule type: protein  
A:Residues: 2-6, 'L', '8-14', 'K', '16', 'X', '18', 'X', '20', 'X', '22-28 <WIT>  
C:Superfamily: rat ribosomal protein L27  
C:Keywords: protein biosynthesis; ribosome  
F:2-136/Product: ribosomal protein L27, cytosolic #status experimental <MAT>

Query Match 76.1%; Score 35; DB 1; Length 136;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKV 7

Db 4 FMKPGKV 10

RESULT 8

A48537  
starch branching enzyme isoform RBE3 - rice  
C:Species: Oryza sativa (rice)  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
A:Accession: A48537  
R:Mizuno, K.; Kawasaki, T.; Shimada, H.; Sato, H.; Kobayashi, E.; Okumura, S.; Arai, Y.  
J. Biol. Chem. 268, 19084-19091, 1993  
A:Title: Alteration of the structural properties of starch components by the lack of an  
A:Reference number: A48537; MUID:93366833  
A:Accession: A48537  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-825 <MIZ>  
A:Cross-references: GB:D16201; NID:g436051; PID:d1004254; PID:g436052  
A:Experimental source: cv. Kinmaze  
A:Note: sequence extracted from NCBI backbone (NCBI:136747, NCBI:136748)  
C:Superfamily: 1,4-alpha-glucan branching enzyme

Query Match 76.1%; Score 35; DB 2; Length 825;

Best Local Similarity 87.5%; Pred. No. 59;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKPGKVKV 9

Db 767 LKPGKVKV 774

RESULT 9

151018  
cobra venom factor precursor - Naja naja  
C:Species: Naja naja  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Feb-1997  
A:Accession: 151018

Query Match 76.1%; Score 35; DB 2; Length 136;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKV 7

Db 151018

RESULT 10

S22288  
ribosomal protein L27, cytosolic - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 14-Nov-1997  
A:Accession: S22288  
R:Lebeau, M.C.; Alvarez-Bolado, G.; Braissant, O.; Wahli, W.; Catsicas, S.  
Nucleic Acids Res. 19, 1337, 1991  
A:Title: Ribosomal protein L27 is identical in chick and rat.  
A:Reference number: S22288; MUID:91232917  
A:Accession: S22288  
A:Molecule type: mRNA  
A:Residues: 1-136 <LEB>  
A:Cross-references: EMBL:X56852; NID:g62980; PID:g62981  
A:Note: the authors translated the codon AAT for residue 127 as Asp  
C:Superfamily: rat ribosomal protein L27  
C:Keywords: protein biosynthesis; ribosome

Query Match 76.1%; Score 35; DB 2; Length 136;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKV 7

Db 151018

RESULT 11

S43505  
ribosomal protein L27, cytosolic - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 14-Nov-1997  
A:Accession: S43505  
R:Gallagher, R.A.; McClean, P.M.; Malik, A.N.  
Biochim. Biophys. Acta 1217, 329-332, 1994  
A:Title: Cloning and nucleotide sequence of a full length cDNA encoding ribosomal pro



phosphomannomutase (EC 5.4.2.8) - Escherichia coli  
 C:Species: Escherichia coli  
 C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 13-Mar-1998  
 A:Accession: I76771; I41252; I62467  
 R: Sugiyama, T.; Kido, N.; Komatsu, T.; Ohta, M.; Jann, K.; Jann, B.; Saeki, A.; Kato, N.  
 Microbiology 140, 59-71, 1994  
 A:Title: Genetic analysis of Escherichia coli O9 rfb: identification and DNA sequence of  
 A:Reference number: I57096; MUID:94214678  
 A:Accession: I76771  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-460 <RES>  
 A:Cross-references: GB:D43637; NID:g598464; PID:g598466  
 R: Jayaratne, P.; Bronner, D.; MacLachlan, P.R.; Dodgson, C.; Kido, N.; Whitfield, C.  
 J. Bacteriol. 176, 3126-3139, 1994  
 A:Title: Cloning and analysis of duplicated rfbM and rfbK genes involved in the formation  
 capsular polysaccharide  
 A:Reference number: I41250; MUID:94252978  
 A:Accession: I41252  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-198, 'V', 200-235, 'E', 237-300, 'N', 302-445, 'A', 447-448, 'E', 450, 'I', 452, 'A', 45  
 A:Cross-references: GB:L27646; NID:g441134; PID:g441137  
 A:Accession: I62467  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-198, 'V', 200-235, 'E', 237-300, 'N', 302-391, 'R', 393-440, 'V', 442-448, 'T', 450-45  
 A:Cross-references: GB:L27632; NID:g441138; PID:g441140  
 C:Keywords: intramolecular transferase; isomerase

Query Match 82.6%; Score 38; DB 2; Length 460;  
 Best Local Similarity 77.8%; Pred. No. 9.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
 |||||:|  
 Db 35 FLKPGKIVV 43

RESULT 4  
 F56146  
 cpsG protein homolog - Klebsiella pneumoniae (fragment)  
 C:Species: Klebsiella pneumoniae  
 C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 09-Sep-1997  
 A:Accession: F56146  
 R: Arakawa, Y.; Wacharotayankun, R.; Nagatsuka, T.; Ito, H.; Kato, N.; Ohta, M.  
 J. Bacteriol. 177, 1788-1796, 1995  
 A:Title: Genomic organization of the Klebsiella pneumoniae cps region responsible for se  
 A:Reference number: A56146; MUID:95204345  
 A:Accession: F56146  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-157 <ARA>  
 A:Cross-references: GB:D21242; NID:g747654; PID:d1005321; PID:g747677

Query Match 82.6%; Score 38; DB 2; Length 157;  
 Best Local Similarity 77.8%; Pred. No. 3.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
 |||||:|  
 Db 35 FLKPGKIVV 43

RESULT 5  
 B35843  
 lipopolysaccharide-binding protein - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 05-Oct-1990 #sequence\_revision 05-Oct-1990 #text\_change 08-Sep-1997  
 A:Accession: B35843; A46553  
 R: Schumann, R.R.; Leong, S.R.; Flagg, G.W.; Gray, P.W.; Wright, S.D.; Mathison, J.C.; T

Science 249, 1429-1431, 1990  
 A:Title: Structure and function of lipopolysaccharide binding protein.  
 A:Reference number: A35843; MUID:90385281  
 A:Accession: B35843  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-482 <SCH>  
 A:Cross-references: GB:M35534; NID:g165467; PID:g165468  
 R: Tobias, P.S.; Soldau, K.; Ulevitch, R.J.  
 J. Exp. Med. 164, 777-793, 1986  
 A:Title: Isolation of a lipopolysaccharide-binding acute phase reactant from rabbit s  
 A:Reference number: A46553  
 A:Accession: A46553  
 A:Molecule type: protein  
 A:Residues: 27-55, 'XG', 58-62, 'F', 64-65 <TOB>  
 C:Superfamily: lipopolysaccharide-binding protein  
 C:Keywords: acute phase

Query Match 78.3%; Score 36; DB 2; Length 482;  
 Best Local Similarity 66.7%; Pred. No. 23;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
 |||||:|  
 Db 402 FLKPGRLQV 410

RESULT 6  
 C3NJ  
 Complement C3 precursor - Naja naja  
 N:Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit  
 C:Species: Naja naja  
 C:Date: 18-Jun-1993 #sequence\_revision 07-Oct-1994 #text\_change 05-Sep-1997  
 C:Accession: A46513  
 R: Fritzinger, D.C.; Petrella, E.C.; Connelly, M.B.; Bredehorst, R.; Vogel, C.W.  
 J. Immunol. 149, 3554-3562, 1992  
 A:Title: Primary structure of cobra complement component C3.  
 A:Reference number: A46513; MUID:93056528  
 A:Accession: A46513  
 A:Molecule type: mRNA  
 A:Residues: 1-1651 <FRI>  
 A:Cross-references: GB:L02365; NID:g213372; PID:g213373  
 A:Note: authors' translation shows Arg-1408 after residue 1438 and, consequently, res  
 A:Note: sequence extracted from NCBI backbone (NCBI:118403) and corrected to corres  
 C:Comment: Complement C3 contains two chains, formed by removal of four residues and  
 alternative complement pathways, releases the C3a anaphylatoxin from the amino end o  
 native-complement-pathway C3/C5 convertase.  
 C:Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.  
 C:Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign  
 e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by  
 C:Comment: The major site of synthesis of this plasma protein is the liver.  
 C:Superfamily: alpha-2-macroglobulin  
 C:Keywords: acute phase; complement alternate pathway; complement pathway; glycoprote  
 F;1-22/Domain: signal sequence #status predicted <SIG>  
 F;23-655/Product: complement C3 and C3b beta chain #status predicted <C3BB>  
 F;23-655,660-1651/Product: complement C3 #status predicted <CC3>  
 F;23-655,739-1651/Product: complement C3b #status predicted <C3B>  
 F;660-1651/Product: complement C3 alpha chain #status predicted <CC3A>  
 F;660-738/Product: C3a anaphylatoxin #status predicted <C3T>  
 F;739-1651/Product: complement C3b alpha' chain #status predicted <C3BA>  
 F;1412-1445/Region: properdin binding  
 F;546-807,615-650,683-710,684-717,697-718,863-1501,1091-1147,1346-1477,1377-1446,1494  
 F;738-739/Cleavage site: Arg-Ser (C3 convertase) #status predicted  
 F;999-1002/Cross-link: thiolester (Cys-Gln) #status predicted

Query Match 76.1%; Score 35; DB 1; Length 1651;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
 |||||:|

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:37 ; Search time 105.14 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-24

Perfect score: 46

Sequence: 1 FLKPGKVKV 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*

1: Pirl.\*

2: Pirl2.\*

3: Pirl3.\*

4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	481	2 A54136	lipopolysaccharide
2	46	100.0	477	2 A35843	lipopolysaccharide
3	38	82.6	460	2 I76771	phosphomannomutase
4	38	82.6	157	2 F56146	cpsg protein homol
5	36	78.3	482	2 B35843	lipopolysaccharide
6	35	76.1	1651	1 C3NJ	complement C3 prec
7	35	76.1	136	1 R5RT27	ribosomal protein
8	35	76.1	825	2 A48537	starch branching e
9	35	76.1	1642	2 I51018	cobra venom factor
10	35	76.1	136	2 S22288	ribosomal protein
11	35	76.1	136	2 S43505	ribosomal protein
12	35	76.1	643	2 E69373	conserved hypothet
13	35	76.1	338	2 C71603	pseudouridine synt
14	34	73.9	886	1 GNLJSP	pol polyprotein -
15	34	73.9	251	2 JC2125	chymase (EC 3.4.21
16	34	73.9	595	2 H69345	translation initia
17	34	73.9	131	2 I38019	melanoma-derived g
18	33	71.7	168	1 CYFGG2	gamma-crystallin I
19	33	71.7	1019	2 F70342	cation efflux syst
20	33	71.7	68	2 S10633	hypothetical prote
21	33	71.7	515	2 E71447	probable PSII D1 p
22	33	71.7	169	2 T01783	gamma-2-crystallin
23	33	71.7	403	2 H70324	hypothetical prote
24	32	69.6	706	1 C42640	kinesin-related pr
25	32	69.6	261	2 D71712	succinate dehydrog
26	32	69.6	226	2 JE0151	myonase (EC 3.4.-
27	32	69.6	244	2 S26042	chymase (EC 3.4.21
28	32	69.6	246	2 B38678	mast cell proteina
29	32	69.6	244	2 A34910	chymase (EC 3.4.21
30	32	69.6	244	2 A46721	ribosomal protein
31	32	69.6	136	2 S46790	ribosomal protein
32	32	69.6	136	2 S69638	phosphate-binding
33	32	69.6	258	2 I64120	translation initia
34	32	69.6	212	2 JN0653	conserved hypothet
35	32	69.6	183	2 F64303	coenzyme F420 hydr
36	32	69.6	627	2 E59504	long-chain-fatty-a
37	32	69.6	645	2 H70173	type III site-spec
38	32	69.6	970	2 S01352	hypothetical prote
39	32	69.6	177	2 S45412	

ALIGNMENTS

RESULT 1

A54136  
lipopolysaccharide-binding protein - human  
N:Alternate names: LBP; LPS-binding protein  
C:Species: Homo sapiens (man)  
C:Date: 11-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 15-Jun-1996  
C:Accession: A54136  
R:Wilde, C.G.; Seilhamer, J.J.; McGrogan, M.; Ashton, N.; Snable, J.L.; Lane, J.C.; L  
J. Biol. Chem. 269, 17411-17416, 1994  
A:Title: Bactericidal/permeability-increasing protein and lipopolysaccharide (LPS)-bi  
A:Reference number: A54136; MUID:94292492  
A:Accession: A54136  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-481 <WIL>  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBIP:149401)  
C:Superfamily: lipopolysaccharide-binding protein

Query Match 100.0%; Score 46; DB 2; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9

|||||

Db 401 FLKPGKVKV 409

RESULT 2

A35843  
lipopolysaccharide-binding protein - human  
C:Species: Homo sapiens (man)  
C:Date: 05-Oct-1990 #sequence\_revision 05-Oct-1990 #text\_change 24-Sep-1998  
C:Accession: A35843  
R:Schumann, R.R.; Leong, S.R.; Flagg, G.W.; Gray, P.W.; Wright, S.D.; Mathison, J.C.  
Science 249, 1429-1431, 1990  
A:Title: Structure and function of lipopolysaccharide binding protein.  
A:Reference number: A35843; MUID:90385281  
A:Accession: A35843  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-477 <SCH>  
A:Cross-references: GB:M35533; NID:gl86965; PID:gl86966  
C:Genetics:  
A:Gene: GDB:LBP  
A:Cross-references: GDB:131571; OMIM:151990  
A:Map position: 20q11.23-20q12  
C:Superfamily: lipopolysaccharide-binding protein

Query Match 100.0%; Score 46; DB 2; Length 477;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9

|||||

Db 397 FLKPGKVKV 405

RESULT 3

I76771

**THIS PAGE BLANK (uspto)**

. Wed Sep 8 09:43:22 1999

US-08-456-112B-24  
; Sequence 24, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-24

Query Match 100.0%; Score 46; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.8e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 FLKPGKVKV 9  
| | | | | | | |  
Db 1 FLKPGKVKV 9

Search completed: September 7, 1999, 23:07:46  
Job time: 1744 sec

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,344
; FILING DATE: 7-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,473
; FILING DATE: 15-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,762
; FILING DATE: 11-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J.
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 93,1133-J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-473-344-98

Query Match 100.0%; Score 46; DB 2; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
Db 401 FLKPGKVKV 409

RESULT 13
US-08-274-303-6
; Sequence 6, Application US/08274303
; Patent No. 5770561
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Arnold H.
; TITLE OF INVENTION: Method For Potentiating BPI Protein Bactericidal
; TITLE OF INVENTION: Activity By Administration Of LBP Proteins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/274,303
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid

US-08-274-303-6

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,344
; FILING DATE: 7-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,473
; FILING DATE: 15-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,762
; FILING DATE: 11-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J.
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 93,1133-J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-473-344-98

Query Match 100.0%; Score 46; DB 2; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
Db 401 FLKPGKVKV 409

RESULT 14
US-08-377-391A-2
; Sequence 2, Application US/08377391A
; Patent No. 5804367
; GENERAL INFORMATION:
; APPLICANT: White, Mark Leslie
; APPLICANT: Carroll, Stephen Fitzhugh
; APPLICANT: Ma, Jeremy Kam-kuen
; TITLE OF INVENTION: METHOD FOR QUANTIFYING LBP IN BODY FLUIDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,391A
; FILING DATE: 24-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/32407
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "rLBP"
; US-08-377-391A-2

Query Match 100.0%; Score 46; DB 2; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
Db 401 FLKPGKVKV 409

RESULT 15
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-205-719-3

Query Match      100.0%; Score 46; DB 1; Length 452;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
Db 376 FLKPGKVKV 384

RESULT 10
US-08-261-660A-4
; Sequence 4, Application US/08261660A
; Patent No. 5731415
; GENERAL INFORMATION:
; APPLICANT: Gazzano-Santoro, Helene
; APPLICANT: Theofan, Georgia
; APPLICANT: Town, Patrick W.
; TITLE OF INVENTION: Lipopolysaccharide Binding Protein Derivatives
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,660A
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/32137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "FLBP"
US-08-261-660A-4

Query Match      100.0%; Score 46; DB 1; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
Db 401 FLKPGKVKV 409

RESULT 11
US-08-209-762-98
; Sequence 98, Application US/08209762
```

```
; Patent No. 5733872
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; TITLE OF INVENTION: Biologically Active Peptides from
; TITLE OF INVENTION: Functional Domains of Bactericidal/Permeability-Increasing
; TITLE OF INVENTION: Protein and Uses Thereof
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,762
; FILING DATE: 11-JAN-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5733872nad, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,1133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "RLBP"
US-08-209-762-98

Query Match      100.0%; Score 46; DB 1; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
Db 401 FLKPGKVKV 409

RESULT 12
US-08-473-344-98
; Sequence 98, Application US/08473344
; Patent No. 5763567
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G
; TITLE OF INVENTION: Biologically Active Peptides from
; TITLE OF INVENTION: Functional Domains of Bactericidal/Permeability-Increasing
; TITLE OF INVENTION: Protein and Uses Thereof
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

GENERAL INFORMATION:  
APPLICANT: Little, Roger G.  
TITLE OF INVENTION: Biologically Active Peptides from  
TITLE OF INVENTION: Functional Domains of Bactericidal/  
TITLE OF INVENTION: Penicillin-Increasing Protein and  
TITLE OF INVENTION: Uses Thereof  
NUMBER OF SEQUENCES: 226  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: Suite 3000, 10 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/306,473A  
FILING DATE: 11-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/209,762  
FILING DATE: 11-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/183,222  
FILING DATE: 14-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: McDonnell, John J.  
REGISTRATION NUMBER: 26,949  
REFERENCE/DOCKET NUMBER: 93,1133-  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 481 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: "rLbp"  
US-08-306-473A-98

Query Match 100.0%; Score 46; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLKPGKVKV 9  
Db 401 FLKPGKVKV 409  
RESULT 8  
US-08-240-328-2  
Sequence 2, Application US/08240328  
Patent No. 5696680  
GENERAL INFORMATION:  
APPLICANT: Ulevitch, Richard  
APPLICANT: Tobias, Peter  
TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING  
TITLE OF INVENTION: ENDOTOXIC ACTIVITY OF LIPOLYSACCHARIDES, ASSAY METHODS AND  
TITLE OF INVENTION: POLYPEPTIDES  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Jubas & Lubitz  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: California  
COUNTRY: USA

ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,328  
FILING DATE: 14-MAY-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: LISA A. TUMARKIN, PH.D.  
REGISTRATION NUMBER: P-38,347  
REFERENCE/DOCKET NUMBER: PD-3575  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 477 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-240-328-2  
Query Match 100.0%; Score 46; DB 1; Length 477;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLKPGKVKV 9  
Db 401 FLKPGKVKV 409  
RESULT 9  
US-08-205-719-3  
Sequence 3, Application US/08205719  
Patent No. 5705398  
GENERAL INFORMATION:  
APPLICANT: Mintz, D. N.  
APPLICANT: Tobias, P. S.  
APPLICANT: Ulevitch, R. J.  
TITLE OF INVENTION: SYSTEM AND METHODS FOR IDENTIFYING  
TITLE OF INVENTION: THERAPEUTIC INHIBITORS OF LPS-MEDIATED SEPSIS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSEE: Patent Counsel  
STREET: 10666 No. 5705398th Torrey Pines Road, TPC 8  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/205,719  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: SCR1201P; TSRI324.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids

REFERENCE/DOCKET NUMBER: 27129/32415  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 481 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "rLBP"  
US-08-372-783-98

Query Match 100.0%; Score 46; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. NO. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKV 9  
| | | | | | | | | |  
Db 401 FLKPGKV 409

RESULT 5  
US-08-372-105-98  
; Sequence 98, Application US/08372105  
; Patent No. 5627153  
; GENERAL INFORMATION:  
; APPLICANT: Little, Roger G.  
; APPLICANT: Lim, Edward  
; APPLICANT: Lambert, Lewis H.  
; APPLICANT: Scannon, Patrick J.  
; TITLE OF INVENTION: Anti-Fungal Materials and Methods  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/372,105  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,540  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,762  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,222  
; FILING DATE: 14-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27129/32415  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 481 amino acids  
; TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "rLBP"  
US-08-372-105-98

Query Match 100.0%; Score 46; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. NO. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKV 9  
| | | | | | | | | |  
Db 401 FLKPGKV 409

RESULT 6  
US-08-097-830E-24  
; Sequence 24, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-24

Query Match 100.0%; Score 46; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. NO. 7.8e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKV 9  
| | | | | | | | | |  
Db 1 FLKPGKV 9

RESULT 7  
US-08-306-473A-98  
; Sequence 98, Application US/08306473A  
; Patent No. 5652332



; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,811  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 27129/31843  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 481 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "rlbp"  
; US-08-186-811-2

Query Match 100.0%; Score 46; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
Db 401 FLKPGKVKV 409

RESULT 3  
US-08-311-611a-98  
; Sequence 98, Application US/08311611a  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611a  
; FILING DATE:

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shaip, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 481 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "rlbp"  
; US-08-311-611a-98

Query Match 100.0%; Score 46; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLKPGKVKV 9  
Db 401 FLKPGKVKV 409

RESULT 4  
US-08-372-783-98  
; Sequence 98, Application US/08372783  
; Patent No. 5578572  
; GENERAL INFORMATION:  
; APPLICANT: Horwitz, Arnold H.  
; APPLICANT: Lambert, Lewis H.  
; APPLICANT: Little, Roger G.  
; TITLE OF INVENTION: Anti-Gram-Positive Bacterial Methods and  
; TITLE OF INVENTION: Materials  
; NUMBER OF SEQUENCES: 237  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/372,783  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,540  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,762  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,222  
; FILING DATE: 14-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:45 ; Search time 82.37 Seconds  
(without alignments)  
1.078 Million cell updates/sec

Title: US-09-124-280A-24  
Perfect score: 46  
Sequence: 1 FLKPGKVKV 9

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2.6/prodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2.6/prodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2.6/prodata/2/iaa/PCTUS9\_COMB.pep:\*  
4: /cgn2.6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	46	100.0	477	1 US-07-847-562-2	Sequence 2, Appli
2	46	100.0	481	1 US-08-186-811-2	Sequence 2, Appli
3	46	100.0	481	1 US-08-311-611A-98	Sequence 98, Appl
4	46	100.0	481	1 US-08-372-783-98	Sequence 98, Appl
5	46	100.0	481	1 US-08-372-105-98	Sequence 98, Appl
6	46	100.0	9	1 US-08-097-830B-24	Sequence 24, Appl
7	46	100.0	481	1 US-08-306-473A-98	Sequence 98, Appl
8	46	100.0	477	1 US-08-240-328-2	Sequence 2, Appli
9	46	100.0	452	1 US-08-205-719-3	Sequence 3, Appli
10	46	100.0	481	1 US-08-261-660A-4	Sequence 4, Appli
11	46	100.0	481	1 US-08-209-762-98	Sequence 98, Appl
12	46	100.0	481	2 US-08-473-344-98	Sequence 98, Appl
13	46	100.0	481	2 US-08-274-303-6	Sequence 6, Appli
14	46	100.0	481	2 US-08-377-391A-2	Sequence 2, Appli
15	46	100.0	9	2 US-08-456-112B-24	Sequence 24, Appl
16	46	100.0	284	2 US-08-215-089-7	Sequence 7, Appli
17	46	100.0	481	2 US-08-215-089-9	Sequence 9, Appli
18	46	100.0	477	2 US-08-485-445A-98	Sequence 98, Appl
19	46	100.0	477	2 US-08-990-849-2	Sequence 2, Appli
20	46	100.0	481	3 PCT-US94-02465-98	Sequence 98, Appl
21	46	100.0	481	3 PCT-US94-06931-4	Sequence 4, Appli
22	46	100.0	481	3 PCT-US94-07834-6	Sequence 6, Appli
23	46	100.0	481	3 PCT-US95-00498-98	Sequence 98, Appl
24	46	100.0	481	3 PCT-US95-00656-98	Sequence 98, Appl
25	46	100.0	284	3 PCT-US95-03384-7	Sequence 7, Appli
26	46	100.0	481	3 PCT-US95-03384-9	Sequence 9, Appli
27	36	78.3	456	1 US-08-205-719-4	Sequence 4, Appli
28	35	76.1	403	1 US-08-118-674-1	Sequence 1, Appli
29	35	76.1	1651	2 US-08-447-411-2	Sequence 2, Appli
30	35	76.1	1642	2 US-08-447-411-45	Sequence 45, Appl
31	35	76.1	1333	2 US-08-447-411-76	Sequence 76, Appl
32	34	73.9	131	2 US-08-578-649-2	Sequence 2, Appli
33	33	71.7	214	3 PCT-US96-08950-2	Sequence 2, Appli
34	33	71.7	214	3 PCT-US96-09127-2	Sequence 2, Appli
35	32	69.6	437	2 US-08-620-605D-4	Sequence 4, Appli
36	31	67.4	64	1 US-08-344-695-17	Sequence 17, Appl
37	31	67.4	510	1 US-08-255-670A-2	Sequence 2, Appli
38	31	67.4	362	1 US-08-255-670A-4	Sequence 4, Appli
39	31	67.4	384	1 US-07-783-706-2	Sequence 2, Appli

40 31 67.4 466 2 US-08-785-066-2 Sequence 2, Appli  
41 31 67.4 384 2 US-08-445-342A-2 Sequence 2, Appli  
42 31 67.4 384 3 PCT-US92-09124-2 Sequence 26, Appli  
43 30 65.2 617 1 US-08-137-614A-26 Sequence 202, App  
44 30 65.2 16 1 US-07-942-245-202 Sequence 203, App  
45 30 65.2 16 1 US-07-942-245-203

ALIGNMENTS

RESULT 1  
US-07-847-562-2  
; Sequence 2, Application US/07847562  
; Patent No. 5310879  
; GENERAL INFORMATION:  
; APPLICANT: Ulevitch, Richard  
; APPLICANT: Tobias, Peter  
; TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC  
; TITLE OF INVENTION: ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND  
; TITLE OF INVENTION: POLYPEPTIDES  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Jubas & Lubitz  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/847,562  
; FILING DATE: 19920302  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wetherell, Jr. Ph.D, John R.  
; REGISTRATION NUMBER: 31,678  
; REFERENCE/DOCKET NUMBER: PD-1859  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 455-5100  
; TELEFAX: (619) 455-5110  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 477 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-847-562-2  
  
Query Match 100.0%; Score 46; DB 1; Length 477;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 FLKPGKVKV 9  
DB 401 FLKPGKVKV 409  
  
RESULT 2  
US-08-186-811-2  
; Sequence 2, Application US/08186811  
; Patent No. 5484705  
; GENERAL INFORMATION:  
; APPLICANT: White, Mark Leslie  
; APPLICANT: Carroll, Stephen Fitzhugh  
; APPLICANT: Ma, Jeremy Ram-Kuen  
; TITLE OF INVENTION: METHOD FOR QUANTIFYING LBP IN BODY FLUIDS  
; NUMBER OF SEQUENCES: 4

CC This is the amino acid sequence of a lipopolysaccharide binding protein  
CC which can be used as a synergist to the actions of a  
CC bactericidal/permeability increasing (BPI) protein (W05852) or peptides  
CC (W05853-W06074) derived from BPI. The BPI peptides can be used to enhance  
CC the treatment of Gram -ve bacterial infections, prophylactically for  
CC patients about to undergo surgery or for disinfection or sterilisation.  
CC The BPI peptides are administered in conjunction with an antibiotic which  
CC both together have synergistic or potentiating effects greater than the  
CC individual peptide or antibiotic. The peptides were screened for  
CC bactericidal activity against E.coli strains J5 and O111:B4 in a radial  
CC diffusion assay.  
SQ Sequence 481 AA:

Query Match 100.0%; Score 46; DB 1; Length 481;

Best Local Similarity 100.0%; Pred. No. 0.38; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
|||||

Db 401 FLKPGKVKV 409

#### RESULT 15

W16796

ID W16796 standard; protein; 487 AA.

AC W16796;

DT 28-JUL-1997 (first entry)

DE Recombinant endotoxin neutralising polypeptide NCY104.

KW Recombinant; endotoxin; neutralisation; RNP; NCY104;

KW lipopolysaccharide; LPS; binding; LBP; bactericidal; chimeric;

KW permeability; increase; BPI; detection; Gram-negative; bacterium;

KW infection; prevention; treatment; disorder; mediation; stimulation;

KW neutrophil; mononuclear cell; inhibition; shock; anemia; anaemia;

KW disseminated intravascular coagulation; thrombocytopenia;

KW thrombocytopenia; adult respiratory distress syndrome; ARDS;

KW renal failure; liver disease; half life; human.

OS Homo sapiens.

OS Synthetic.

PH Key Location/Qualifiers

FT peptide

FT 1..31 /label= sig\_peptide

FT region 32..230

FT /note= "human BPI residues 1-199"

FT region 231..487

FT /note= "human LBP residues 200-456"

PN W09634873-A1.

PD 07-NOV-1996.

PF 01-MAY-1996; U06134.

PR 01-MAY-1995; US-431517.

PA (INCY-) INCYTE PHARM INC.

PI Marra MN, Scott RW;

PI WPI; 96-506090/50

PT Lipo:polysaccharide binding and endotoxin neutralising polypeptide

PT - useful in diagnosis, prevention and treatment of Gram-negative

PT bacterial infection and associated disorders and conditions

PS Example 1: Page 1: 147pp; English.

CC The present sequence is the recombinant endotoxin neutralising

CC polypeptide (RNP) B(1-199)(200-456) (NCY104), which comprises

CC the human bactericidal/permeability increasing protein (hBPI)

CC signal peptide, hBPI residues 1-199 and human lipopolysaccharide

CC (LPS) binding protein (hLBP) residues 200-256.

CC The RNP, which selectively and specifically binds LPS and has

CC endotoxin neutralising activity, can be used to detect a site of

CC Gram negative bacterial infection, and prevent or treat endotoxin

CC related disorders, preferably where LPS mediated stimulation of

CC neutrophils and mononuclear cells is inhibited, e.g. shock,

CC disseminated intravascular coagulation, anaemia, thrombocytopenia,

CC adult respiratory distress syndrome, renal failure, liver disease

CC and conditions associated with Gram negative bacterial infection.

CC The RNP has an enhanced serum half life relative to naturally

CC occurring LBP, and binds LPS without triggering a significant,

CC undesirable immune response.

CC N.B. Sequence not given in specification, but produced using the  
CC wild type hLBP and hBPI sequences given in figures 4 and 3.  
SQ Sequence 487 AA:

Query Match 100.0%; Score 46; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
|||||

Db 407 FLKPGKVKV 415

Search completed: September 7, 1999, 22:49:56

Job time: 7877 sec

Wed Sep 8 09:43:21 1999

us-09-124-280a-24.rag

```

QY 1 FLKPGKVKV 9
DB 385 FLKPGKVKV 393

RESULT 12
R81246
ID R81246 standard; Protein: 481 AA.
AC R81246;
DT 15-MAR-1996 (first entry)
DE Recombinant lipopolysaccharide binding protein.
KW Bactericidal/permeability increasing protein; BPI; granule; mammalian;
KW polymorphonuclear neutrophil; anti-bacterial; fungus; infection;
KW antifungal; fluconazole; amphotericin B; Candida albicans; sterilise;
KW lipopolysaccharide binding protein; sterilisation; medical instrument.
OS Synthetic.
FH Key
FT peptide 1. .25 Location/Qualifiers
FT peptide 26. .481 /note= "signal peptide"
FT peptide /note= "mature peptide"
FT W09519179-A1.
PD 20-JUL-1995.
PF 13-JAN-1994; US-183222.
PR 14-JAN-1994; US-209762.
PR 11-JUL-1994; US-273540.
PR 11-JUL-1994; US-273540.
PA (XOMA ) XOMA CORP.
PI Little RG, Lambert LH, Lim E, Scannon PJ;
PI WPI: 95-263713/34.
DR N-PSDB: Q99815.
DR Treating fungal infection with bactericidal permeability increasing
PT protein or deriv. - esp. for control of systemic Candida albicans
PT infection or for use in in vitro sterilisation
PS Disclosure; Page 84-86; 133pp; English.
CC The amino acid sequence of a lipopolysaccharide binding protein (LBP).
CC The protein is used to enhance the antifungal activity of a
CC bactericidal/permeability increasing (BPI) protein-derived peptide
CC (see R81083-4, R81088-81244 and R81248-R81308 for examples of peptides).
CC The BPI protein can be isolated from the granules of mammalian
CC polymorphonuclear neutrophils (PMN). The protein has antibacterial
CC activity associated with 3 functional domains: I (R81085), II (R81086)
CC and III (R81087) present in N-terminal region of the BPI holoprotein.
CC The peptides are used to treat fungal infections together with other
CC antifungal cpds e.g. fluconazole or amphotericin B. The peptides can be
CC used to treat fungal infection, esp. Candida albicans. They are also
CC useful for killing or inhibiting fungi in vitro e.g. for sterilising
CC medical instruments.
CC Sequence 481 AA;

Query Match 100.0%; Score 46; DB 1; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
DB 401 FLKPGKVKV 409

RESULT 13
R80994
ID R80994 standard; Protein: 481 AA.
AC R80994;
DT 23-MAY-1996 (first entry)
DE Lipopolysaccharide binding protein (rLBP25).
KW Bactericidal/permeability increasing peptide; BPI; heparin; binding;
KW neutralisation; lipopolysaccharide; LPS; bactericidal activity;
KW treatment; neutralise endotoxin; inhibit angiogenesis;
KW inhibit tumour formation; proliferation.
OS Homo sapiens.

Query Match 100.0%; Score 46; DB 1; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
DB 401 FLKPGKVKV 409

RESULT 14
W05851
ID W05851 standard; Protein: 481 AA.
AC W05851;
DT 18-DEC-1996 (first entry)
DE Recombinant lipopolysaccharide binding protein.
KW Lipopolysaccharide binding protein; synergist; BPI; enhancement;
KW bactericidal/permeability increasing protein; potentiation; surgery;
KW Gram negative; bacterial infections; disinfection; sterilisation;
KW antibiotic; E.coli.
OS Synthetic.
FH Key
FT peptide 1. .25 Location/Qualifiers
FT peptide 26. .481 /note= "signal peptide"
FT protein /note= "mature recombinant lipopolysaccharide binding
protein"
FT US5232388-A.
PD 04-JUN-1996.
PF 22-SEP-1993; US-125651.
PR 22-SEP-1993; US-125651.
PR 11-JUL-1994; US-273401.
PR 22-SEP-1994; US-311611.
PA (XOMA ) XOMA CORP.
PI Cohen J, Kung AHC, Lambert LH, Little RG;
PI WPI: 96-285780/29.
DR N-PSDB: T39395.
DR Compsn. for treating Gram negative bacterial infection - contg.
PT antibiotic and bactericidal-permeability increasing protein as
PT synergist
PS Disclosure; Column 129-134; 138pp; English.

```

```

FH Key
FT peptide 1. .26 Location/Qualifiers
FT peptide /note= "signal peptide"
FT protein 27. .481 /note= "mature protein"
FT W09519372-A1.
PD 20-JUL-1995.
PF 15-SEP-1994; U10427.
PR 14-JAN-1994; US-183222.
PR 11-MAR-1994; US-209762.
PA (XOMA ) XOMA CORP.
PI Little RG;
PI WPI: 95-263828/34.
DR N-PSDB: T00950.
DR New peptide(s) based on bactericidal/permeability-increasing protein
PT - having heparin binding and neutralisation, LPS binding and
PT neutralisation and antimicrobial activities
PS Disclosure; Page 126-128; 275pp; English.
CC Recombinant lipopolysaccharide binding protein (rLBP25) is encoded
CC by T00950. LBP protein products are useful in conjunction with BPI
CC peptides. Esp. pref. LBP derivs. are amino acids of the amino terminus of
CC characterised by the first 197 amino acids of the amino acid
CC rLBP shown in R80994. BPI (bactericidal permeability-increasing)
CC peptides (R80996-81081 and R82553-372) each have an amino acid
CC sequence that is deriv. of a BPI functional domain (or a variant)
CC having at least one of the biological activities of BPI, such as
CC heparin binding or neutralisation; lipopolysaccharide (LPS) binding
CC or neutralisation or bactericidal activity. The BPI peptides are
CC based on the amino-terminal portion of BPI, esp. functional domains
CC I, II, and III (BPI residues 17-45, 65-99 and 142-169 resp.).
CC Sequence 481 AA;

Query Match 100.0%; Score 46; DB 1; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9
DB 401 FLKPGKVKV 409

RESULT 14
W05851
ID W05851 standard; Protein: 481 AA.
AC W05851;
DT 18-DEC-1996 (first entry)
DE Recombinant lipopolysaccharide binding protein.
KW Lipopolysaccharide binding protein; synergist; BPI; enhancement;
KW bactericidal/permeability increasing protein; potentiation; surgery;
KW Gram negative; bacterial infections; disinfection; sterilisation;
KW antibiotic; E.coli.
OS Synthetic.
FH Key
FT peptide 1. .25 Location/Qualifiers
FT peptide 26. .481 /note= "signal peptide"
FT protein /note= "mature recombinant lipopolysaccharide binding
protein"
FT US5232388-A.
PD 04-JUN-1996.
PF 22-SEP-1993; US-125651.
PR 22-SEP-1993; US-125651.
PR 11-JUL-1994; US-273401.
PR 22-SEP-1994; US-311611.
PA (XOMA ) XOMA CORP.
PI Cohen J, Kung AHC, Lambert LH, Little RG;
PI WPI: 96-285780/29.
DR N-PSDB: T39395.
DR Compsn. for treating Gram negative bacterial infection - contg.
PT antibiotic and bactericidal-permeability increasing protein as
PT synergist
PS Disclosure; Column 129-134; 138pp; English.

```

PI Carroll SF, White ML;  
 DR WPI; 95-269544/35.  
 DR N-PSDB; Q92931.  
 PT Sandwich immunoassay for detection of lipo:polysaccharide in body  
 PT fluids - useful in diagnostic and prognostic methods for evaluating  
 PT gram-negative sepsis  
 PS Disclosure; Page 17-18; 35pp; English.  
 CC The AA and nt sequence of recombinant LBP are set out in co-owned  
 CC and copending U.S. Patent Application Serial No. 08/029,510 filed  
 CC June 17, 1993, and shown in Q92931 & R76601. A recombinant LBP amino-  
 CC terminal fragment is characterized by the AA sequence of the first  
 CC 197 AAs of the amino-terminus of LBP, as set out in Q92932 & R76602,  
 CC the production of which is described in co-owned and copending U.S.  
 CC Patent Application Serial No. 08/079,510 filed June 12 1993, the  
 CC disclosure of which is incorporated. LBP is a 60 kD glycoprotein  
 CC synthesised in the liver. Schumann et al., Science, 249:1429 (1990)  
 CC disclose the AA sequences and encoding cDNA of both human and rabbit  
 CC LBP. LBP has been characterised as an 'acute phase protein'.  
 SQ Sequence 481 AA;

Query Match 100.0%; Score 46; DB 1; Length 481;

Best Local Similarity 100.0%; Pred. No. 0.38; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

QY 1 FLKPGKVKV 9  
 |||||

DB 401 FLKPGKVKV 409

## RESULT 9

ID R80014  
 AC R80014 standard; Protein; 481 AA.  
 DT 15-JAN-1996 (first entry)  
 DE Human LBP.  
 KW Lipopolysaccharide binding protein; LBP; sepsis; antiseptic;  
 KW CD14.  
 OS Homo sapiens.  
 PN W09525117-A1.  
 PD 21-SEP-1995.  
 PF 15-MAR-1995; U03384.  
 PR (SCRI) SCRIPPS RES INST.  
 PA Han J, Tobias PS, Ulevitch RJ;  
 PI WPI; 95-336937/43.  
 DR N-PSDB; T04519.  
 PT Isolated lipo:polysaccharide binding polypeptide (LPS BP) useful for  
 PT treating sepsis - prevents the LPS BP complex from transferring LPS  
 PT to CD14  
 PS Disclosure; Fig. 5A-5C; 64pp; English.  
 CC Novel LBPs comprising amino acids 1-197 (given in R80012) or  
 CC 198-481 (R80013) of human LBP have antiseptic properties. The  
 CC LBPs are derived from full-length human LBP (R80014) encoded  
 CC by the DNA given in T04519.  
 SQ Sequence 481 AA;

Query Match 100.0%; Score 46; DB 1; Length 481;

Best Local Similarity 100.0%; Pred. No. 0.38; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

QY 1 FLKPGKVKV 9  
 |||||

DB 401 FLKPGKVKV 409

## RESULT 10

ID R80013  
 AC R80013 standard; protein; 284 AA.  
 DT 15-JAN-1996 (first entry)  
 DE Human LBP(198-481).

KW Lipopolysaccharide binding protein; LBP; sepsis; antiseptic;  
 KW CD14.  
 OS Homo sapiens.  
 PN W09525117-A1.  
 PD 21-SEP-1995.  
 PF 15-MAR-1995; U03384.  
 PR 15-MAR-1994; US-215089.  
 PA (SCRI) SCRIPPS RES INST.  
 PI Han J, Tobias PS, Ulevitch RJ;  
 DR WPI; 95-336937/43.  
 DR N-PSDB; T04518.  
 PT Isolated lipo:polysaccharide binding polypeptide (LPS BP) useful for  
 PT treating sepsis - prevents the LPS BP complex from transferring LPS  
 PT to CD14  
 PS Claim 1; Page 36-37; 64pp; English.  
 CC A truncated LBP, comprising amino acids 198-481 of human LBP, does  
 CC not bind to lipopolysaccharide (LPS), but does bind to CD14, and can  
 CC inhibit the interaction of LPS:LBP complex with CD14. It is used  
 CC as an antiseptic or to raise antibodies.  
 SQ Sequence 284 AA;

Query Match 100.0%; Score 46; DB 1; Length 284;

Best Local Similarity 100.0%; Pred. No. 0.23; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

QY 1 FLKPGKVKV 9  
 |||||

DB 204 FLKPGKVKV 212

## RESULT 11

ID R76242  
 AC R76242 standard; Protein; 465 AA.  
 DT 01-FEB-1996 (first entry)  
 DE Recombinant LBP.  
 KW BPI peptide; bacterial permeability-increasing peptide; bactericidal;  
 KW therapeutic effectiveness; antibiotic; concurrent administration;  
 KW reverse resistance; gram-negative bacteria.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..25 /note= "signal peptide"  
 FT protein 26..465 /note= "mature rLBP"  
 FT W09508344-A1.  
 PD 30-MAR-1995.  
 PR 22-SEP-1994; U11225.  
 PR 22-SEP-1993; US-125651.  
 PR 11-JUL-1994; US-273401.  
 PA (XOMA) XOMA CORP.  
 PI Cohen J, Kung AHC, Lambert LH, Little RG;  
 DR WPI; 95-161465/21.  
 DR N-PSDB; Q92639.  
 PT BPI protein and an antibiotic in a medicament - for treatment of  
 PT gram-negative bacterial infection  
 PS Disclosure; Page 172-174; 259pp; English.  
 CC The DNA encodes a recombinant LBP (sic). The specification discloses  
 CC a recombinant BPI (bactericidal-permeability increasing protein),  
 CC fragments of which (R76244-458) were screened for bactericidal effects  
 CC on E. coli strains J5 and 0111:B4 in a radial diffusion assay. BPI  
 CC peptides which retain antibacterial activity are expected to improve  
 CC the therapeutic effectiveness of antibiotics when concurrently  
 CC administered. Concurrent administration of BPI protein products and  
 CC antibiotics is shown to reverse resistance of a variety of gram-negative  
 CC organisms to antibiotics.  
 SQ Sequence 465 AA;

Query Match 100.0%; Score 46; DB 1; Length 465;

Best Local Similarity 100.0%; Pred. No. 0.36; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides. 9 AA;  
SQ Sequence 9 AA;

Query Match 100.0%; Score 46; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
| | | | | | | | | |  
DB 1 FLKPGKVKV 9

RESULT 7  
R68922  
ID R68922 standard; Protein; 481 AA.  
AC R68922;  
DT 11-SEP-1995 (first entry)  
DE Lipopolysaccharide binding protein (rLBP).  
KW Lipopolysaccharide binding protein; rLBP; LPS;  
KW gram-negative bacterial infections; treatment.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT peptide 1..25  
FT /label= sig\_peptide  
PN W09500641-A.  
PD 05-JAN-1995.  
PF 17-JUN-1994; U06931.  
PR 17-JUN-1993; US-079510.  
PA (XOMA ) XOMA CORP.  
PI Gazzano-santoro H, Theofan G, Trown PW;  
DR WPI: 95-052078/07.  
DR N-PSDB: Q80826.  
PT Lipopolysaccharide binding protein deriv. and hybrid protein  
PT binds to lipo-polysaccharide - lacks CD14-mediated  
PT immuno-stimulatory properties, used to treat gram-negative  
PT bacterial infections and associated conditions  
PS Example 2; Pages 63-65; 114pp; English.  
CC Q80826 encodes R68922 recombinant lipopolysaccharide (LPS) binding  
CC protein rLBP. The protein R68915 derived from R68922 lacks  
CC CD14-mediated immunostimulatory properties, and can therefore be used  
CC to treat gram-negative bacterial infections and associated  
CC conditions.  
SQ Sequence 481 AA;

Query Match 100.0%; Score 46; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
| | | | | | | | | |  
DB 401 FLKPGKVKV 409

RESULT 8  
R76601  
ID R76601 standard; Protein; 481 AA.  
AC R76601;  
DT 22-DEC-1995 (first entry)  
DE Recombinant human lipopolysaccharide binding protein (rLBP).  
KW Lipopolysaccharide; outer membrane; gram-negative bacteria;  
KW endotoxin; assay; lipopolysaccharide binding protein; LBP;  
KW glycoprotein; acute phase protein.  
OS Homo sapiens.  
PN W09520163-A.  
PD 27-JUL-1995.  
PR 24-JAN-1995; U00982.  
PR 24-JAN-1994; US-186811.  
PA (XOMA ) XOMA CORP.

DB 401 FLKPGKVKV 409  
| | | | | | | | | |

RESULT 5  
R68000  
ID R68000 standard; Protein; 481 AA.  
AC R68000;  
DT 15-AUG-1995 (first entry)  
DE rLBP  
KW Bactericidal/permeability-increasing protein; BPI; rBPI; rBPI50;  
KW lipopolysaccharide binding protein; LBP; rLBP25; rLBP;  
KW Gram-negative bacterium; infection; antiseptic.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT peptide 1..25  
FT /label= Sig\_peptide  
PN W09502414-A.  
PD 26-JAN-1995.  
PF 13-JUL-1994; U07834.  
PR 14-JUL-1993; US-093201.  
PR 11-JUL-1994; US-274303.  
PA (XOMA ) XOMA CORP.  
PI Horwitz A;  
DR WPI: 95-067161/09.  
DR N-PSDB: Q81447.  
PT Method for treating gram negative bacterial infection - comprises  
PT administering lipopolysaccharide binding protein (LBP) prod. and  
PT bactericidal/permeability-increasing (BPI) protein prod.  
PS Disclosure: Page 43-44; 76pp; English.  
CC Gram-negative bacterial infections are treated by co-administration  
CC of BPI protein and LBP. A preferred LBP is the 197-amino acid  
CC N-terminal fragment (given in R67999) of recombinant LBP (rLBP,  
CC R67800).  
SQ Sequence 481 AA;

Query Match 100.0%; Score 46; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
| | | | | | | | | |  
DB 401 FLKPGKVKV 409

RESULT 6  
R71795  
ID R71795 standard; peptide; 9 AA.  
AC R71795;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porto M;  
DR WPI: 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 26; Page 22; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which  
CC include units of formula: (A)n, where A is the cationic amino acid Lys  
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
CC Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock

FT US245013-A. /label= mature\_LBP  
 PD 14-SEP-1993.  
 PF 30-APR-1985; 728833.  
 PR 30-APR-1985; US-728833.  
 PR 30-DEC-1986; US-006710.  
 PR 16-JUN-1989; US-367454.  
 PA (ULEV/) ULEVITCH R.  
 PI Tobias P, Ulevitch R;  
 DR WPI: 93-302704/38.  
 DR N-PSDB; Q47905.  
 DE New glycoprotein which binds Gram negative lipo:polysaccharide -  
 PT isolated from acute phase sera, useful for determ. of endotoxin  
 in body fluid  
 PS Claim 1; Fig 19; 41pp; English.  
 CC A human liver cDNA library was prepared using Poly-A+ RNA isolated  
 CC from liver biopsies of trauma patients. The library was transferred  
 CC to nitrocellulose and screened twice with a probe based on the  
 CC N-terminal sequence of rabbit LBP. The coding sequence obtained from  
 CC positive clones encodes a glycoprotein which binds to a lipopoly-  
 CC saccharide secreted by a gram negative bacterium. The glycoprotein  
 CC retards in vitro binding of LPS to high density lipoprotein present  
 CC in normal serum.  
 CC Sequence 477 AA;  
 SQ

Query Match 100.0%; Score 46; DB 1; Length 477;  
 Best Local Similarity 100.0%; Pred. No. 0.37;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
 DB 397 FLKPGKVKV 405  
 |||||

RESULT 3  
 ID R53459 standard; Protein: 477 AA.  
 AC R53459;  
 DT 08-DEC-1994 (first entry)  
 DE Human lipopolysaccharide binding protein.  
 KW hLBP; human lipopolysaccharide binding protein; detection;  
 OS Gram negative bacterial infection; ss.  
 FH Key  
 FT peptide 1..25  
 FT protein /label= signal peptide  
 FT protein 26..477  
 FT protein /label= lipopolysaccharide binding protein  
 FT protein /note= "mature protein"

US5310879-A.  
 PD 10-MAY-1994.  
 PF 30-APR-1985; 728833.  
 PR 30-APR-1985; US-728833.  
 PR 30-DEC-1986; US-006710.  
 PR 16-JUN-1989; US-367454.  
 PR 02-MAR-1992; US-847562.  
 PA (SCRI ) SCRIPPS RES INST.  
 PI Tobias PS, Ulevitch RJ;  
 DR WPI: 94-150499/18.  
 DR N-PSDB; Q63542.  
 DE Antibodies immunoreactive with lapine lipopolysaccharide binding  
 PT protein - useful for determ. of presence and amt. of human and  
 PT rabbit LBP.  
 PS Disclosure; Fig 19; 44pp; English.  
 CC R53459 shows a human lipopolysaccharide binding protein (LBP).  
 CC LBP is a glycoprotein that binds lipopolysaccharide that can be  
 CC present in the bloodstream of an individual susceptible to infection  
 CC by LPS-secreting gram negative bacteria.  
 CC Sequence 477 AA;  
 SQ

Query Match 100.0%; Score 46; DB 1; Length 477;  
 Best Local Similarity 100.0%; Pred. No. 0.37;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 0.37;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9  
 DB 401 FLKPGKVKV 409  
 |||||

RESULT 4  
 ID R62501 standard; Protein: 481 AA.  
 AC R62501;  
 DT 03-MAY-1995 (first entry)  
 DE LBP amino terminal fragment, designated rLBP25.  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW contraception; malignant; tumour cell; inflammatory disease; T-cell;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipoarabinomannan; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 FH Key  
 FT peptide 1..25  
 FT protein /note= "Signal peptide"  
 FT protein 26..481  
 FT protein /note= "LBP N-terminal fragment"

WO9420532-A.  
 PD 15-SEP-1994.  
 PF 11-MAR-1994; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.  
 PA (XOMA ) XOMA CORP.  
 PI Little RG;  
 DR WPI: 94-302964/37.  
 DR N-PSDB; Q74302.  
 DE New human bactericidal permeability increasing peptides - derived  
 PT from the functional domains of BPI and having BPI activities such  
 PT as bactericidal activity  
 PS Disclosure; Page 151-154; 254pp; English.  
 CC This sequence represents the N-terminal fragment of lipopolysaccharide  
 CC binding protein (LBP) which has a molecular weight of approx. 25 kD.  
 CC This protein is designated rLBP25 and may be used in conjunction with  
 CC the sequences given in R63682-750, R62087-100 and R62491-500 which are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC protein (BPI). BPI derived peptides such as these, may be used as  
 CC heparin binding agents, for neutralising the anti-coagulant effect of  
 CC heparin, for inhibiting angiogenesis, eg, associated with ocular  
 CC retinopathy, for inhibiting endothelial cell proliferation, for  
 CC contraception, for inhibiting malignant tumour cell proliferation,  
 CC for treating a chronic inflammatory disease state, eg, rheumatoid  
 CC arthritis, and for treating gram-negative bacterial infection. The  
 CC peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipoarabinomannan in the circulation,  
 CC eg, a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipoarabinomannan or for treating a disease associated with Helicobacter  
 CC infection, eg, gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg, gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 481 AA;  
 SQ

Query Match 100.0%; Score 46; DB 1; Length 481;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLKPGKVKV 9

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.  
Cobra partial CVP2  
Potato starch bran

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:49:56 ; Search time 135.78 Seconds  
(without alignments)  
1.570 Million cell updates/sec

Title: US-09-124-280A-24  
Perfect score: 46  
Sequence: 1 FLKPGKVKV 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	481	1 R34140	LBP-beta. New form
2	46	100.0	477	1 R41660	Human gram-negativ
3	46	100.0	477	1 R53459	Human lipopolysacc
4	46	100.0	481	1 R2501	LBP amino terminal
5	46	100.0	481	1 R68000	rLBP. Method for t
6	46	100.0	9	1 R71795	Peptide neutralisi
7	46	100.0	481	1 R68922	Lipopolysaccharide
8	46	100.0	481	1 R76601	Recombinant human
9	46	100.0	481	1 R80014	Human LBP. Isolate
10	46	100.0	284	1 R80013	Human LBP(198-481)
11	46	100.0	465	1 R7242	Recombinant LBP. B
12	46	100.0	481	1 R81246	Recombinant lipopo
13	46	100.0	481	1 R80994	Lipopolysaccharide
14	46	100.0	481	1 W05851	Recombinant lipopo
15	46	100.0	487	1 W16796	Recombinant endoto
16	46	100.0	481	1 W16829	Recombinant endoto
17	46	100.0	481	1 W18830	Recombinant endoto
18	46	100.0	481	1 W18831	Recombinant endoto
19	46	100.0	257	1 W16835	Recombinant endoto
20	46	100.0	481	1 W16836	Recombinant endoto
21	46	100.0	481	1 W16837	Recombinant endoto
22	46	100.0	481	1 W16838	Recombinant endoto
23	46	100.0	481	1 W18839	Recombinant endoto
24	46	100.0	481	1 W18840	Recombinant endoto
25	46	100.0	481	1 W19010	Recombinant endoto
26	46	100.0	481	1 W19011	Recombinant endoto
27	46	100.0	482	1 W16819	Recombinant endoto
28	46	100.0	482	1 W16820	Recombinant endoto
29	46	100.0	480	1 W18821	Recombinant endoto
30	46	100.0	480	1 W16822	Recombinant endoto
31	46	100.0	9	1 W21512	Antibiotic potenti
32	46	100.0	481	1 W40812	Human lipopolysacc
33	46	100.0	452	1 W41694	Human LPS-binding
34	46	100.0	481	1 W40139	Human recombinant
35	46	100.0	481	1 W63305	Lipopolysaccharide
36	46	100.0	481	1 W77093	Human lipopolysacc
37	46	100.0	481	1 W78356	Human gram-negati
38	36	78.3	482	1 R41659	Lipine polysacchar
39	36	78.3	482	1 R53461	Rabbit lipopolysac
40	36	78.3	482	1 W40813	Rabbit lipopolysac
41	36	78.3	456	1 W41695	Rabbit LPS-binding
42	36	78.3	482	1 W78355	Rabbit lipopolysac
43	35	76.1	825	1 W60811	Rice starch branch

ALIGNMENTS

RESULT 1

R34140  
ID R34140 standard; Protein; 481 AA.  
AC R34140;  
DT 20-JUL-1993 (first entry)  
DE LBP-beta.  
KW Liposaccharide binding protein; immunostimulant; antibody;  
KW immunoassay; LBP-alpha; recombinant.  
OS Synthetic.  
FH key Location/Qualifiers  
FT region 25..481  
FT peptide /note= "mature protein"  
FT 1..24 /note= "signal peptide"  
PN WO9306228-A.  
PD 01-APR-1993.  
PF 28-SEP-1992; U08298.  
PR 26-SEP-1991; US-765660.  
PA (INCY-) INCYTE PHARM INC.  
PI Delegeane AM, Seilhamer JJ;  
DR WPI: 93-117548/14.  
DR N-PSDB; Q39305.  
PT New form of lipo-saccharide binding protein, LBP-beta - useful as immunostimulant, and antibodies are used to detect LBP-beta or related proteins  
PS Claim 1; Fig 2; 19pp; English.  
CC PCR primers were used to amplify liposaccharide binding protein beta encoding DNA from a commercially available cDNA prep. derived from human liver RNA by standard PCR. The amplified prods. were digested with the corresp. restriction enzymes and isolated by PAGE. The fragments were ligated into pMamNeo, a mammalian expression vector, and cloned into E. coli. DNA from resulting clones was analysed and sequenced. The sequence is similar to that of Schumann et al., Science (1990) 249: 1429-1431 but also contains an additional sequence of 12 nucleotides in the coding region beginning after base 826 of Genbank sequence 35533. The LBP-beta may be used to stimulate the immune system and permits fine tuning of therapeutic protocols designed to enhance immune function. Antibodies specifically reactive with LBP-beta may be used in immunoassay systems for detection of LBP-beta or related proteins. Antibodies which recognise epitopes related to the novel portion of LBP-beta which distinguish it from the previously disclosed human LBP are useful (there are 7 amino acid substns. in LBP-beta relative to LBP-alpha).  
SQ Sequence 481 AA;

Query Match 100.0%; Score 46; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. No. 0.38; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 FLKPGKVKV 9  
|||||  
DB 401 FLKPGKVKV 409

RESULT 2

R41660  
ID R41660 standard; Protein; 477 AA.  
AC R41660;  
DT 18-MAR-1994 (first entry)  
DE Human gram-negative lipopolysaccharide binding protein.  
DE Gram negative bacteria; LPS; high density lipoprotein;  
KW human; glycoprotein; gastrointestinal tract; endotoxaemia.  
OS Homo sapiens.  
FH key Location/Qualifiers  
FT Protein 26..477



**THIS PAGE BLANK (USPTO)**

Db 804 IKQSFVK 811  
:::|||||

## RESULT 14

Q17645 PRELIMINARY; PRT: 1283 AA.  
AC Q17645;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-WAY-1999 (TReMBLrel. 10, Last annotation update)  
DE C05A9.1.1 PROTEIN.  
GN C05A9.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KERSHAW J.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans.";  
RL Nature 368:32-38(1994).  
DR EMBL; 270265; CAA94202.1; -.  
DR PFAM; PF00664; ABC\_membrane; 2.  
DR PFAM; PF00005; ABC\_tran; 2.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
KW ATP-binding; Transport.  
SQ SEQUENCE 1283 AA; 144323 MW; 1AB726E6 CRC32;

Query Match 70.5%; Score 31; DB 5; Length 1283;

Best Local Similarity 62.5%; Pred. No. 3.7e+02; Mismatches 2; Indels 0; Gaps 0;

QY 1 KVRKSFFK 8

I:|I|I|I|

Db 928 KIRNSYFK 935

## RESULT 15

Q21368 PRELIMINARY; PRT: 267 AA.  
AC Q21368;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-JAN-1999 (TReMBLrel. 09, Last annotation update)  
DE K08H2.3 PROTEIN.  
GN K08H2.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GAJADSTY S.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,

RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans.";  
RL Nature 368:32-38(1994).  
DR EMBL; 270210; CAA94150.1; -.  
SQ SEQUENCE 267 AA; 31200 MW; F24D80C4 CRC32;

Query Match 70.5%; Score 31; DB 5; Length 267;

Best Local Similarity 62.5%; Pred. No. 86; Mismatches 2; Indels 0; Gaps 0;

QY 1 KVRKSFFK 8

I:|I|I|I|

Db 96 KIRKKEFE 103

Search completed: September 7, 1999, 22:47:31

Job time: 7951 sec

RESULT 10  
O67862  
ID O67862 PRELIMINARY; PRT; 524 AA.  
AC O67862;  
DT 01-AUG-1998 (TReMBLrel. 07, Created)  
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
DE 2-ISOPROPYLMALATE SYNTHASE.  
GN LEUA2.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF5;  
RX MEDLINE: 98196666.  
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
aeolicus.";  
RL Nature 392:353-358(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF5;  
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: A8000772; AAC07824.1; -;  
DR PFAM: PF00682; HMGL-like; 1.  
DR PROSITE: PS00815; AIPM\_HOMOCIT\_SYNTH\_1; 1.  
DR PROSITE: PS00816; AIPM\_HOMOCIT\_SYNTH\_2; 1.  
SQ SEQUENCE 524 AA; 58562 MW; D1FD2D75 CRC32;

Query Match 70.5%; Score 31; DB 2; Length 524;  
Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSF 7  
DB 253 KVRKDF 259  
|||||

RESULT 11  
O43962  
ID Q43962 PRELIMINARY; PRT; 362 AA.  
AC Q43962;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
DE REPB.  
GN REPB.  
OS Clostridium butyricum.  
OG Plasmid pCB101.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
OC Clostridium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCIB 7423;  
RA BREHM J.K., PENNOCK A., YOUNG M., OULTRAM J.D., MINTON N.P.;  
RL Plasmid 0:0-0(0).  
DR EMBL: X62684; CAA44562.1; -;  
DR PFAM: PF01446; Rep; 1.  
KW Plasmid.  
SQ SEQUENCE 362 AA; 43052 MW; 8BF2B2F1 CRC32;

Query Match 70.5%; Score 31; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKSEFK 8  
DB 353 RKSEFK 358  
|||||

RESULT 12  
O75165  
ID O75165 PRELIMINARY; PRT; 1021 AA.  
AC O75165;  
DT 01-NOV-1998 (TReMBLrel. 08, Created)  
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
DE KIAA0678 PROTEIN (FRAGMENT).  
GN KIAA0678.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE: 98403880.  
RA ISHIKAWA K., NAGASE T., SUYAMA M., MIYAJIMA N., TANAKA A., KOTANI H.,  
RA NOMURA N., OHARA O.;  
RT "Prediction of the coding sequences of unidentified human genes. X.  
The complete sequences of 100 new cDNA clones from brain which can  
code for large proteins in vitro.";  
RL DNA Res. 5:169-176(1998).  
DR EMBL: AB014578; BAA31653.1; -;  
DR PFAM: PF00226; DnaJ; 1.  
FT NON\_TER 1  
SQ SEQUENCE 1021 AA; 114534 MW; B5ED52B CRC32;

Query Match 70.5%; Score 31; DB 4; Length 1021;  
Best Local Similarity 44.4%; Pred. No. 3e+02;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9  
DB 97 KIRKAYFRL 105  
|||||

RESULT 13  
O94821  
ID O94821 PRELIMINARY; PRT; 945 AA.  
AC O94821;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
DE KIAA0713 PROTEIN (FRAGMENT).  
GN KIAA0713.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA NAGASE T., ISHIKAWA K., SUYAMA M., KIKUNO R., MIYAJIMA N., TANAKA A.,  
RA KOTANI H., NOMURA N., OHARA O.;  
RT "Prediction of the coding sequences of unidentified human genes. XI.  
The complete sequences of 100 new cDNA clones from brain which code  
for large proteins in vitro.";  
RL DNA Res. 5:277-286(1998).  
DR EMBL: AB018256; BAA34433.1; -;  
FT NON\_TER 1  
SQ SEQUENCE 945 AA; 108276 MW; D7E7892F CRC32;

Query Match 70.5%; Score 31; DB 4; Length 945;  
Best Local Similarity 62.5%; Pred. No. 2.8e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRKSEFKV 9

```

ID O67616 PRELIMINARY; PRT; 385 AA.
AC O67616;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE HYPOTHETICAL 45.8 KD PROTEIN.
GN AQ.1723.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE000753; AAC07586.1; -.
DR PFAM; PF01142; UPF0024; 1.
DR PROSITE; PS01268; UPF0024; 1.
KW Hypothetical protein.
SQ SEQUENCE 385 AA; 45774 MW; EC864FA9 CRC32;

Query Match 72.7%; Score 32; DB 2; Length 385;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRKSFFK 8
Db 339 VRKAFFK 345

RESULT 7
O77393 PRELIMINARY; PRT; 2437 AA.
AC O77393;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MAL3P6.2 PROTEIN.
GN MAL3P6.2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-3D7;
RA MURPHY L., LAWSON D., BARRELL B.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98551; CAB1149.1; -.
SQ SEQUENCE 2437 AA; 295348 MW; E6280382 CRC32;

Query Match 72.7%; Score 32; DB 5; Length 2437;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFFK 8
Db 1152 KVRMEFK 1159

RESULT 8
Q39049 PRELIMINARY; PRT; 1381 AA.
AC Q39049;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
DE MAGNESIUM CHELATASE SUBUNIT.
GN CHLH.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C24;
RA GIBSON L., MARRISON J., LEECH R., JENSEN P., BASSHAM D., GIBSON M.,
RA HUNTER C.;
RL Plant Physiol. 11:61-71(1996).
DR EMBL; Z68495; CAA92802.1; -.
DR MENDEL; 17280; ARATH; 2627; mnl7280.
SQ SEQUENCE 1381 AA; 153508 MW; 5AB299B6 CRC32;

Query Match 72.7%; Score 32; DB 10; Length 1381;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KVRKSFFK 9
Db 39 KPAKSFFK 47

RESULT 9
O67640 PRELIMINARY; PRT; 678 AA.
AC O67640;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HYPOTHETICAL 79.4 KD PROTEIN.
GN AQ.1761.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE000755; AAC07610.1; -.
KW Hypothetical protein.
SQ SEQUENCE 678 AA; 79356 MW; E1AB53B9 CRC32;

Query Match 70.5%; Score 31; DB 2; Length 678;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFFK 9
Db 633 KVSEFFK 641

```

```

SQ SEQUENCE 1461 AA: 166226 MW: 26C3FEED CRC32:

Query Match 79.5%; Score 35; DB 3; Length 1461;
Best Local Similarity 87.5%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFFK 8
   III IIII
Db 1215 KVRNSFFK 1222

RESULT 3
O25728 ID O25728 PRELIMINARY; PRT; 290 AA.
AC O25728;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE CONSERVED HYPOTHETICAL SECRETED PROTEIN.
GN HP1098.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695;
RX MEDLINE: 97394467.
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATHEY L., WALLIN E.,
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA VENTER J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
DR EMBL: AE000616; AAD08141.1;
DR TIGR: HP1098;
KW Hypothetical protein.
SQ SEQUENCE 290 AA: 31594 MW: 2C92CBC2 CRC32:

Query Match 75.0%; Score 33; DB 2; Length 290;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRKSFFKV 9
   I:IIII:I
Db 5 VKKSFFRV 12

RESULT 4
P95069 ID P95069 PRELIMINARY; PRT; 220 AA.
AC P95069;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE RPSE.
GN RPSE.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA DEVLIN K., CHURCHER C.M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]

SQ SEQUENCE 1461 AA: 166226 MW: 26C3FEED CRC32:

Query Match 75.0%; Score 33; DB 3; Length 362;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFK 8
   :::IIII
Db 313 RIKKSFFK 320

RESULT 6
O67616
```

```

RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA BARRELL B.G., RAJANDREAM M.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE: 96181548.
RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RT "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37Rv, and comparison with Mycobacterium
RT leprae.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).
DR EMBL: 284395; CAB06445.1;
DR PFAM: PF00333; Ribosomal_S5; 1.
DR PROSITE: PS00585; RIBOSOMAL_S5; 1.
SQ SEQUENCE 220 AA: 22888 MW: 964B0EF6 CRC32:

Query Match 75.0%; Score 33; DB 2; Length 220;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9
   : IIIII:I
Db 97 EARKSFFRV 105

RESULT 5
O92406 ID O92406 PRELIMINARY; PRT; 362 AA.
AC O92406;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE P43-ACMNPV ORF39.
GN P43.
OS Bombyx mori nuclear polyhedrosis virus (BmNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T3;
RX MEDLINE: 97329351.
RA KAMITA S.G., MAEDA S.;
RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx
RT mori nuclear polyhedrosis virus and fine-mapping of a region involved
RT in host range expansion.";
RL Gene 190:173-179(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=T3;
RA GOMI S., MAJIMA K., MAEDA S.;
RT "Sequence analysis of the genome of Bombyx mori
RT nucleopolyhedrovirus.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: L33180; AAC63714.1;
SQ SEQUENCE 362 AA: 43424 MW: 1379543C CRC32:

Query Match 75.0%; Score 33; DB 12; Length 362;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFK 8
   :::IIII
Db 313 RIKKSFFK 320

RESULT 6
O67616
```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:29 ; Search time 160.2 Seconds  
(without alignments)  
3.458 Million cell updates/sec

Title: US-09-124-280A-23

Perfect score: 44

Sequence: 1 KVRKSFVKV 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- SPTREMBL10.\*  
1: sp-archaea.\*  
2: sp-bacteria.\*  
3: sp-fungi.\*  
4: sp-human.\*  
5: sp-invertebrate.\*  
6: sp-mammal.\*  
7: sp-mhc.\*  
8: sp-organelle.\*  
9: sp-phage.\*  
10: sp-plant.\*  
11: sp-rodent.\*  
12: sp-virus.\*  
13: sp-vertebrate.\*  
14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	79.5	215	2	Q45214 borrelia he
2	35	79.5	1451	3	O94532 schizosacch
3	33	75.0	290	2	O25728 helicobacte
4	33	75.0	220	2	P95069 mycobacteri
5	33	75.0	362	12	O92406 bombyx mori
6	32	72.7	385	2	O67616 aquifex aeo
7	32	72.7	2437	5	O77393 plasmodium
8	32	72.7	1381	10	Q39049 arabidopsis
9	31	70.5	678	2	O67640 aquifex aeo
10	31	70.5	524	2	O67862 aquifex aeo
11	31	70.5	362	2	O45962 clostridium
12	31	70.5	1021	4	O75165 homo sapien
13	31	70.5	945	4	O94821 homo sapien
14	31	70.5	1283	5	O17645 caenorhabdi
15	31	70.5	267	5	Q21368 caenorhabdi
16	31	70.5	1158	5	Q22655 caenorhabdi
17	31	70.5	1114	5	Q22656 caenorhabdi
18	31	70.5	297	10	O82163 arabidopsis
19	31	70.5	725	10	P92949 arabidopsis
20	31	70.5	1765	11	O88457 rattus norv
21	30	68.2	98	1	Q30139 archaeoglob
22	30	68.2	353	2	O25770 helicobacte
23	30	68.2	437	2	O51145 borrelia bu
24	30	68.2	232	2	O66786 aquifex aeo
25	30	68.2	353	2	Q92K59 helicobacte
26	30	68.2	345	5	P91030 caenorhabdi
27	30	68.2	996	5	O01425 caenorhabdi
28	30	68.2	344	5	O01788 caenorhabdi
29	30	68.2	381	5	Q94198 caenorhabdi

30	68.2	691	5	O44741	caenorhabdi
31	68.2	566	5	O16580	caenorhabdi
32	68.2	802	5	O45413	caenorhabdi
33	68.2	1280	6	O46605	canis famil
34	68.2	230	10	O22736	arabidopsis
35	68.2	99	10	O82499	arabidopsis
36	65.9	241	1	O57895	pyrococcus
37	65.9	215	2	P96566	amycolatops
38	65.9	408	2	P96561	amycolatops
39	65.9	373	2	P96619	bacillus su
40	65.9	452	2	O59917	treponema h
41	65.9	280	2	O51409	borrelia bu
42	65.9	355	2	O51697	borrelia bu
43	65.9	726	2	P72664	synechocyst
44	65.9	284	2	O66464	aquifex aeo
45	65.9	297	10	Q92P26	arabidopsis

ALIGNMENTS

RESULT 1  
Q45214 PRELIMINARY; PRT; 215 AA.  
ID O45214  
AC Q45214;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE VARIABLE OUTER MEMBRANE PROTEIN.  
GN VMP6.  
OS Borrelia hermsii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H51 SEROTYPE 6;  
RX MEDLINE: 94373822.  
RA RESPEFO B.I.; BARBOUR A.G.;  
RT "Antigen diversity in the bacterium B. hermsii through "somatic"  
RT mutations in rearranged vmp genes.";  
RL Cell 78:867-876(1994).  
DR EMBL: L33898; AAA59224.1; .  
DR PFAM: PF01441; Lipoprotein\_6; 1.  
SQ SEQUENCE 215 AA; 22728 MW; 8903AAAF2 CRC32;

Query Match 79.5%; Score 35; DB 2; Length 215;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFVKV 9  
||| |||  
DB 172 KVRKSLFKL 180

RESULT 2  
O94532 PRELIMINARY; PRT; 1461 AA.  
ID O94532  
AC O94532;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE PUTATIVE INVOLVEMENT IN CYTOSKELETAL ORGANISATION.  
GN SPCC895.05.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972H-;  
RA WOOD V., RAJANDREAM M.A., BARRELL B.G., LUCAS M., GAILLARDIN C.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL035247; CAA22841.1; .

**THIS PAGE BLANK (USPTO)**

entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
RESULT 14
RL1L CAEEL STANDARD; PRT; 196 AA.
ID RL1L CAEEL
AC Q94300;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROBABLE 60S RIBOSOMAL PROTEIN L11.
GN T2F3.4.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SEGERMENTEA; RHABDITIA; RHABDITIDA.
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WU X.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: PROBABLY ASSOCIATES WITH 5S RNA.
CC -!- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
EMBL; U64844; G1658341; -
WORMPEP; T2F3.4; C613968.
DR PROSITE; PS00358; RIBOSOMAL_L5; 1.
DR PFAM; PF00281; L5; 1.
DR PFAM; PF00673; L5.C; 1.
DR RIBOSOMAL PROTEIN.
KW RIBOSOMAL PROTEIN.
SQ SEQUENCE 196 AA; 22714 MW; CC94C9F6 CRC32;
```

Query Match 70.5%; Score 31; DB 1; Length 196;  
Best Local Similarity 62.5%; Pred. No. 12; Indels 0; Gaps 0;  
Matches 5; Conservative 3; Mismatches 0;

QY 1 KVRKSFFK 8  
|||:|:|:  
DB 187 KVKKNFFR 194

```
RESULT 15
RL1L ODOSI STANDARD; PRT; 130 AA.
ID RL1L ODOSI
AC P49499;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S11.
GN RPS11.
OS ODONTELLA SINENSIS.
OG CHLOROPLAST.
OC EUKARYOTA; STRAMENOPILES; BACILLARIOPHYTA; COSCINODISCOPHYCEAE;
OC BIDDULPHIOPHYCIDAE; EUPODISCALES; EUPODISCACEAE; ODONTELLA.
RN [1]
RP SEQUENCE FROM N.A.
RA KOWALLIK K.V., STOEBE B., SCHAFFRAN I., KROTH-PANCIC P., FREIER U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis".
RL PLANT MOL. BIOL. REP. 13:336-342(1995).
CC -!- SIMILARITY: BELONGS TO THE S11P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
EMBL; Z67753; E211821; -
DR PROSITE; PS00054; RIBOSOMAL_S11; 1.
DR PFAM; PF00411; S11; 1.
DR MENDEL; 4128; ODOsi; rps11; 1.
KW RIBOSOMAL PROTEIN; CHLOROPLAST.
SQ SEQUENCE 130 AA; 13772 MW; 6D1CA336 CRC32;
```

Query Match 70.5%; Score 31; DB 1; Length 130;  
Best Local Similarity 75.0%; Pred. No. 7.8; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 1;

QY 1 KVRKSFFK 8  
|:|:|:|:  
DB 4 KIRKSTFK 11

Search completed: September 7, 1999, 23:59:03  
Job time: 515 sec



```

Db 461 KVKENYFKV 469
||||:||||
RESULT 12
PIGR_RABIT
ID PIGR_RABIT STANDARD; PRT; 773 AA.
AC P01832;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE POLYMERIC-IMMUNOGLOBULIN RECEPTOR PRECURSOR (PLGR) [CONTAINS:
DE SECRETORY COMPONENT].
GN PIGR.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84142246.
RA MOSTOV K.E., FRIEDLANDER M., BLOBEL G.;
RT "The receptor for transepithelial transport of IgA and IgM contains
RT multiple immunoglobulin-like domains.";
RL NATURE 308:37-43(1984).
RN [2]
RP SEQUENCE OF 87-114 AND 410-428.
RX MEDLINE; 88228032.
RA FRUTIGER S., HUGHES G.J., HANLY W.C., JATON J.-C.;
RT "Rabbit secretory components of different allotypes vary in their
RT carbohydrate content and their sites of N-linked glycosylation.";
RL J. BIOL. CHEM. 263:8120-8125(1988).
CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC -!- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF ALLOTYPIC T62.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC 5 V-LIKE DOMAINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X00412; G1596; -.
DR EMBL; K01291; G165106; -.
DR PIR; A02111; ORRBG
DR PIR; A28077; A28077.
DR PFAM; PF00047; 19; 5.
KW IMMUNOGLOBULIN FOLD; REPEAT; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL;
KW POLYMORPHISM.
FT SIGNAL 1 18
FT CHAIN 19 773 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 615 SECRETORY COMPONENT.
FT DOMAIN 19 647 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 648 670 POTENTIAL.
FT DOMAIN 671 773 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 136 IG-LIKE V-TYPE DOMAIN 1.
FT DOMAIN 137 243 IG-LIKE V-TYPE DOMAIN 2.
FT DOMAIN 244 350 IG-LIKE V-TYPE DOMAIN 3.
FT DOMAIN 351 456 IG-LIKE V-TYPE DOMAIN 4.
FT DOMAIN 457 558 IG-LIKE V-TYPE DOMAIN 5.
FT DISULFID 46 115 POTENTIAL.
FT DISULFID 155 225 POTENTIAL.
FT DISULFID 260 324 POTENTIAL.
FT DISULFID 369 438 POTENTIAL.
FT DISULFID 478 538 POTENTIAL.
-----
FT CARBOHYD 88 88 IN ALLOTYPIC T61.
FT CARBOHYD 108 108 IN ALLOTYPIC T62 (PARTIAL) AND T63.
FT CARBOHYD 418 418
FT VARIANT 88 88 K -> N (IN ALLOTYPIC T61).
FT VARIANT 94 94 D -> E (IN ALLOTYPIC T61).
FT VARIANT 101 108 TVDQLTON -> YLNRLSOS (IN ALLOTYPIC T61).
FT VARIANT 110 110 S -> T (IN ALLOTYPIC T63).
SQ SEQUENCE 773 AA; 83886 MW; 79840DIF CRC32;
-----
Query Match 70.5%; Score 31; DB 1; Length 773;
Best Local Similarity 55.6%; Pred. No. 51;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVRKSPFKV 9
DB 163 QLKSPFKV 171
:::||||:|
-----
RESULT 13
RK32_VICFA
ID RK32_VICFA STANDARD; PRT; 48 AA.
AC P15820;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L32.
GN RPL32.
OS VICIA FABA (BROAD BEAN).
OG CHLOROPLAST.
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC FABALES; FABACEAE; PAPILIONOIDEAE; VICIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90206803.
RA HERDENBERGER F., PILLAY D.T.N., STEINMETZ A.;
RT "Sequence of the trnH gene and the inverted repeat structure deletion
RT site of the broad bean chloroplast genome.";
RL NUCLEIC ACIDS RES. 18:1297-1297(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89208884.
RA HERDENBERGER F., WEIL J.H., STEINMETZ A.;
RT "Organization and nucleotide sequence of the broad bean chloroplast
RT genes trnL-UAG, ndhF and two unidentified open reading frames.";
RL CURR. GENET. 14:609-615(1988).
CC -!- SIMILARITY: BELONGS TO THE L32P FAMILY OF RIBOSOMAL PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X51471; G12389; -.
DR PIR; S08495; RSVF32.
DR PIR; S08444; S08444.
DR MENDEL; 5435; VICFA; rpl32:1.
KW RIBOSOMAL PROTEIN; CHLOROPLAST.
SQ SEQUENCE 48 AA; 5407 MW; 03503C8B CRC32;
-----
Query Match 70.5%; Score 31; DB 1; Length 48;
Best Local Similarity 62.5%; Pred. No. 2.7;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVRKSPFK 8
DB 14 KIRKFNWK 21
|:|:|:|:|

```

Wed Sep 8 09:43:19 1999

RT lipopolysaccharide binding protein. Evidence for extrahepatic

RT expression."

RL J. IMMUNOL. 153:743-752(1994).

CC -|- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL

CC LIPOPOLYSACCHARIDES (LPS), A GLYCOPOLIPID PRESENT IN THE OUTER

CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS

CC TO INTERACT WITH THE CD14 RECEPTOR.

CC -|- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC EMBL: L32132; G510715; "

CC PROSITE; PS00400; LBP\_BPI\_CETP; 1.

CC PFAM; PF01273; LBP\_BPI\_CETP; 1.

CC HSP; P17213; LBP1.

CC LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.

CC SIGNAL 1 25 POTENTIAL.

CC CHAIN 26 481 LIPOPOLYSACCHARIDE-BINDING PROTEIN.

CC CARBOHYD 300 300 POTENTIAL.

CC CARBOHYD 355 355 POTENTIAL.

CC SEQUENCE 481 AA; 53600 MW; 0BF55E7D CRC32;

SQ

Query Match 70.5%; Score 31; DB 1; Length 481;

Best Local Similarity 66.7%; Pred. No. 31;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFVK 9

Db 117 KVRSFVKL 125

RESULT 10

ID PALY\_RHOBB STANDARD; PRT; 713 AA.

AC P10248; (REL. 10, CREATED)

AT 01-MAR-1989 (REL. 13, LAST SEQUENCE UPDATE)

DT 01-JAN-1990 (REL. 32, LAST ANNOTATION UPDATE)

DE 01-NOV-1995 (REL. 4.3.1.5).

GN PHENYLALANINE AMMONIA-LYASE

OS RHODOTORULA RUBRA.

OC EUKARYOTA; FUNGI; BASIDIOMYCOTA; UREDINIOMYCETES;

OC MITOSPORIC UREDINIOMYCETES; SPOROBOLOMYCETACEAE; RHODOTORULA.

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN-NRRL Y-15597;

RX MEDLINE; 89083582.

RA FILPULA D., STRAUSBERG R.L., VASLET C.A., SYKES A., LEVY A.;

RT "Nucleotide sequence of gene for phenylalanine ammonia-lyase from

RT Rhodotorula rubra.";

RL NUCLEIC ACIDS RES. 16:11381-11381(1988).

CC [2]

CC DETAILS.

CC STRAIN-NRRL Y-15597;

RX MEDLINE; 89083583.

RA FILPULA D., STRAUSBERG R.L., VASLET C.A., SYKES A., LEVY A.;

RT "cDNA and genomic cloning of yeast phenylalanine ammonia-lyase genes

RT reveal genomic intron deletions";

RL NUCLEIC ACIDS RES. 16:11382-11382(1988).

CC -|- CATALYTIC ACTIVITY: L-PHENYLALANINE = TRANS-CINNAMATE + NH(3).

CC -|- COFACTOR: HAS A DEHYDROLANINE (DHA) IN THE ACTIVE SITE.

CC -|- PATHWAY: KEY ENZYME OF PHENYLPROPANOID METABOLISM.

CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).

CC -|- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC EMBL: X13094; G295942; "

CC EMBL: X13095; E10865; ALT\_SEQ.

CC PIR; S01999; S01999.

CC PROSITE; PS00488; PAL\_HISTIDASE; 1.

CC PFAM; PF00221; PAL; 1.

CC LYASE; PHENYLPROPANOID METABOLISM.

CC ACT SITE 218 218 BY SIMILARITY.

CC SEQUENCE 713 AA; 76000 MW; 2FECBAC0 CRC32;

SQ

Query Match 70.5%; Score 31; DB 1; Length 713;

Best Local Similarity 75.0%; Pred. No. 47;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFVK 8

Db 573 KVRKSIK 580

RESULT 11

ID PFL\_CLOPA STANDARD; PRT; 740 AA.

AC Q46266; (REL. 35, CREATED)

AT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE FORMATE ACETYLTRANSFERASE (EC 2.3.1.54) (PYRUVATE FORMATE-LYASE).

GN PFL.

OS CLOSTRIDIUM PASTEURIANUM.

OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;

OC CLOSTRIDIUM.

CC [1]

CC SEQUENCE FROM N.A.

CC MEDLINE; 96218720.

RX WEIDNER G., SAWERS G.;

RT "Molecular characterization of the genes encoding pyruvate

RT formate-lyase and its activating enzyme of Clostridium

RT pasteurianum";

RL J. BACTERIOL. 178:2440-2444(1996).

CC -|- CATALYTIC ACTIVITY: ACETYL-COA + FORMATE = COA + PYRUVATE.

CC -|- PATHWAY: GLUCOSE METABOLISM (NONOXIDATIVE CONVERSION).

CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).

CC -|- SIMILARITY: STRONG, TO OTHER PYRUVATE FORMATE-LYASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC EMBL: X93463; G1072361; "

CC PROSITE; PS00850; GLY\_RADICAL; 1.

CC PFAM; PF01228; GLY\_RADICAL; 1.

CC TRANSFERASE; ACYLTRANSFERASE; GLUCOSE METABOLISM; ORGANIC RADICAL.

FT MOD\_RES 715 715 FREE RADICAL (BY SIMILARITY).

CC SEQUENCE 740 AA; 83216 MW; 4CE5E8C5 CRC32;

SQ

Query Match 70.5%; Score 31; DB 1; Length 740;

Best Local Similarity 55.6%; Pred. No. 48;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFVK 9

```

DR EMBL: 298756; E337977; -.
DR PROSITE: PS00585; RIBOSOMAL_S5; 1.
DR PFAM: PF00333; S5; 1.
DR HSP: P02357; 1PKP.
KW RIBOSOMAL PROTEIN.
SQ SEQUENCE 217 AA; 22614 MW; DDDDC767 CRC32;

Query Match          75.0%; Score 33; DB 1; Length 217;
Best Local Similarity 66.7%; Pred. No. 5.1;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9
   : |||||:
Db 93 EARKSFRV 101

RESULT 7
UCRO_HUMAN          STANDARD;          PRT;          92 AA.
ID UCR0_HUMAN
AC Q14949;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 13-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX UBIQUINONE-BINDING PROTEIN
DE QP-C (EC 1.10.2.2) (UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 9.5 KD
DE PROTEIN) (COMPLEX III SUBUNIT VII).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]
RN SEQUENCE FROM N.A.
RP FUJIMURA T., KAWAI A., SHIMIZU F., SHINOMIYA K., HIRANO H., OKUNO S.,
RA OZAKI K., KATAGIRI T., TAKEDA S., KUGA Y., SHIMADA Y., NAGATA M.,
RA TAKACHI A., WATANABE T., HORIE M., NAKAMURA Y., TAKAHASHI E.,
RA HIRAI Y.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: THIS IS A COMPONENT OF THE UBIQUINOL-CYTOCHROME C
CC REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH
CC IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THIS SUBUNIT,
CC TOGETHER WITH CYTOCHROME B, BINDS TO UBIQUINONE.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C = Q +
CC 2 FERROCYTOCHROME C.
CC -1- SUBUNIT: BCL COMPLEX CONTAINS 10 SUBUNITS; 3 RESPIRATORY
CC SUBUNITS, 2 CORE PROTEINS AND 5 LOW-MOLECULAR WEIGHT PROTEINS.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -1- SIMILARITY: TO YEAST SUBUNIT 8 (QCR8).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: D50369; D1024200; -.
DR MITOCHONDRION; INNER MEMBRANE; ELECTRON TRANSPORT; RESPIRATORY CHAIN;
KW OXIDOREDUCTASE.
FT INIT_MET 0
SQ SEQUENCE 92 AA; 11185 MW; F220250C CRC32;

Query Match          75.0%; Score 33; DB 1; Length 92;
Best Local Similarity 55.6%; Pred. No. 2;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9
   :||:||||:
Db 40 RIRESEFRV 48

```

```

RESULT 8
KAD5_SCHPO          STANDARD;          PRT;          566 AA.
ID KAD5_SCHPO
AC Q09831;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROBABLE SERINE/THREONINE-PROTEIN KINASE C4G8.05 (EC 2.7.1.1.-).
GN SPAC4G8.05.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-972;
RC BADCOCK K., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: TO YEAST PROTEIN KINASE KIN82.
CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: Z56276; G1022350; -.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM: PF00069; pkinase; 1.
DR HSP: Q03450; 1A06.
KW HYPOTHETICAL PROTEIN; TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE;
KW ATP-BINDING.
FT DOMAIN 195 485 PROTEIN KINASE.
FT NP_BIND 201 209 ATP (BY SIMILARITY).
FT BINDING 224 224 ATP (BY SIMILARITY).
FT ACT_SITE 320 320 BY SIMILARITY.
SQ SEQUENCE 566 AA; 63482 MW; D50F780B CRC32;

Query Match          70.5%; Score 31; DB 1; Length 566;
Best Local Similarity 71.4%; Pred. No. 36;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRKSFFK 8
   :|||||
Db 116 IKKSFFK 122

RESULT 9
LBP_RAT            STANDARD;          PRT;          481 AA.
ID LBP_RAT
AC Q63313;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).
GN LBP.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RC MEDLINE; 9429804.
RX SU G.L., FREESWICK P.D., GELLER D.A., WANG Q., SHAPIRO R.A., WAN Y.H.,
RA BILLIAR T.R., TWEARDY D.J., SIMMONS R.L., WANG S.C.;
RT "Molecular cloning, characterization, and tissue distribution of rat

```

Wed Sep 8 09:43:19 1999

```

DR  PIR: A45684; A45684.
SQ  SEQUENCE 363 AA; 43490 MW; 0E93F359 CRC32;

Query Match      77.3%; Score 34; DB 1; Length 363;
Best Local Similarity 75.0%; Pred. No. 5.3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY  1 KVRKSEFFK 8
    :|:|||||
Db  313 RYKSEFFK 320

RESULT 4
Y129_BORBU      STANDARD;      PRT;      249 AA.
AC  051155;
DT  15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT  15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE  HYPOTHETICAL PROTEIN BB0129.
GN  BB0129.
OS  BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
OC  BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 35210 / B31;
RX  MEDLINE; 98065943.
RA  FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA  LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R.G., HICKEY E.K., GWINN M.,
RA  DOUGHERTY B., TOMB J.F., FLEISCHMANN R.D., RICHARDSON D., HANSON M.,
RA  PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., WEIDMAN J.,
RA  VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., BOWMAN C.,
RA  UTTERBACK T., WATHEY L., McDONALD L., ARTIACH P., BOWMAN C.,
RA  GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA  SMITH H.O., VENTER J.C.;
RT  "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT  burgdorferi."
RL  NATURE 390:580-586(1997).
CC  -!- SIMILARITY: BELONGS TO FAMILY 1 OF PSEUDOURIDINE SYNTHASES.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: AEO00124; G2688006; -
CC  TIGR: BB0129; -
CC  PROSITE: PS01149; PSI_1; 1.
KW  HYPOTHETICAL PROTEIN.
SQ  SEQUENCE 249 AA; 28949 MW; 263994A72 CRC32;

Query Match      77.3%; Score 34; DB 1; Length 249;
Best Local Similarity 66.7%; Pred. No. 3.6;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY  1 KVRKSEFFK 9
    :|:|||||
Db  163 KVRKSEFFK 171

RESULT 5
OPPF_MYCPN
ID  OPPE_MYCPN      STANDARD;      PRT;      851 AA.
AC  P75551;
DT  01-NOV-1997 (REL. 35, CREATED)
DT  01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT  01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE  OLIGOPEPTIDE TRANSPORT ATP-BINDING PROTEIN OPPE.

GN  OPPE.
OS  MYCOPLASMA PNEUMONIAE.
OC  BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;
OC  MYCOPLASMATACEAE; MYCOPLASMA.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 29342 / M129;
RX  MEDLINE; 97105885.
RA  HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
RA  HERRMANN R.;
RT  "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT  pneumoniae."
RT  NUCLEIC ACIDS RES. 24:4420-4449(1996).
RL  FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC  FOR OLIGOPEPTIDES. MAY BE RESPONSIBLE FOR ENERGY COUPLING TO THE
CC  TRANSPORT SYSTEM (BY SIMILARITY).
CC  -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (POTENTIAL).
CC  -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC  (ABC TRANSPORTERS).
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
EMBL: AEO00058; G1674315; -
DR  PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR  PFAM; PF00005; ABC_tran; 1.
KW  PEPTIDE TRANSPORT; MEMBRANE; ATP-BINDING.
FT  NP_BIND 48 55
FT  SEQUENCE 851 AA; 98611 MW; D2888B67 CRC32;
SQ  SEQUENCE 851 AA; 98611 MW; D2888B67 CRC32;

Query Match      75.0%; Score 33; DB 1; Length 851;
Best Local Similarity 87.5%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 KVRKSEFFK 8
    :|:|||||
Db  22 KVRKSEFFK 29

RESULT 6
RS5_MYCLE
ID  RS5_MYCLE      STANDARD;      PRT;      217 AA.
AC  O33000;
DT  15-JUL-1998 (REL. 36, CREATED)
DT  15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT  15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE  30S RIBOSOMAL PROTEIN S5.
GN  RPSE OR MLCB2492.21.
OS  MYCOBACTERIUM LEPRAE.
OS  BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
OC  ACTINOMYCETALES; CORYNEBACTERIACEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
RN  [1]
RP  SEQUENCE FROM N.A.
RA  COLE S.T., FLESSELLES B., HONORE N.;
RA  SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RL  FUNCTION: PROTEIN S5 IS IMPORTANT IN THE ASSEMBLY AND FUNCTION OF
CC  THE 30S RIBOSOMAL SUBUNIT (BY SIMILARITY).
CC  -!- SIMILARITY: BELONGS TO THE S5P FAMILY OF RIBOSOMAL PROTEINS.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----

```

```

DR EMBL; X98665; E264344; JOINED.
DR EMBL; X98666; E264344; JOINED.
DR EMBL; X98667; E264344; JOINED.
DR EMBL; X98668; E264344; JOINED.
DR EMBL; AF013512; G2653817; JOINED.
DR EMBL; AF013500; G2653817; JOINED.
DR EMBL; AF013501; G2653817; JOINED.
DR EMBL; AF013502; G2653817; JOINED.
DR EMBL; AF013503; G2653817; JOINED.
DR EMBL; AF013504; G2653817; JOINED.
DR EMBL; AF013505; G2653817; JOINED.
DR EMBL; AF013506; G2653817; JOINED.
DR EMBL; AF013507; G2653817; JOINED.
DR EMBL; AF013508; G2653817; JOINED.
DR EMBL; AF013509; G2653817; JOINED.
DR EMBL; AF013510; G2653817; JOINED.
DR EMBL; AF013511; G2653817; JOINED.
DR EMBL; L42172; G807652; -.
DR PIR; A35843; A35843.
DR MIM; 151990; -.
DR PROSITE; PS00400; LBP_BPI_CETP; 1.
DR PFAM; PF01273; LBP_BPI_CETP; 1.
DR HSP; P17213; LBP1.
KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 481 LIPOLYSACCHARIDE-BINDING PROTEIN.
FT CARBOHYD 300 300 POTENTIAL.
FT CARBOHYD 355 355 POTENTIAL.
FT CARBOHYD 386 386 POTENTIAL.
FT CARBOHYD 394 394 POTENTIAL.
FT CONFLICT 82 82 N -> K (IN REF. 3).
FT CONFLICT 128 128 S -> F (IN REF. 3).
FT CONFLICT 154 157 VTAS -> GYCL (IN REF. 1).
FT CONFLICT 174 174 L -> S (IN REF. 1).
FT CONFLICT 257 257 R -> S (IN REF. 3).
FT CONFLICT 266 270 VMSLP -> A (IN REF. 1).
FT CONFLICT 369 369 L -> H (IN REF. 3).
FT CONFLICT 436 436 L -> F (IN REF. 3).
FT SEQUENCE 481 AA; 53349 MW; 9F3D0008 CRC32;

Query Match 93.28; Score 41; DB 1; Length 481;
Best Local Similarity 88.98; Pred. No. 0.24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSEFKV 9
Db 117 KVRKSEFKL 125

RESULT 2
LBP_MOUSE
ID LBP_MOUSE STANDARD; PRT; 481 AA.
AC Q61805;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).
GN LBP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
[1]
RC SEQUENCE FROM N.A.
RC STRAIN=BALE/C.
RA MEDLINE; 97289150.
RA LENGACHER S., JONGENEEL C.V., LE ROY D., LEE J.D., KRAVCHENKO V.,
RA ULEVITCH R.J., GLAUSER M.P., HEUMANN D.;
RT "Reactivity of murine and human recombinant LPS-binding protein (LBP)
RT within LPS and gram negative bacteria."
RL J. INFLAMM. 47:165-172(1995).
CC -!- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL
CC LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER

```

---

```

CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS
CC TO INTERACT WITH THE CD14 RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X99347; E253895; -.
CC MGD; MGI:1098776; LBP.
CC PROSITE; PS00400; LBP_BPI_CETP; 1.
CC PFAM; PF01273; LBP_BPI_CETP; 1.
CC HSP; P17213; LBP1.
KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 481 LIPOPOLYSACCHARIDE-BINDING PROTEIN.
FT CARBOHYD 300 300 POTENTIAL.
FT CARBOHYD 355 355 POTENTIAL.
FT SEQUENCE 481 AA; 53312 MW; 973CFF58 CRC32;

Query Match 79.5%; Score 35; DB 1; Length 481;
Best Local Similarity 77.8%; Pred. No. 4.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KVRKSEFKV 9
Db 117 KVRKSEFKL 125

RESULT 3
VP43_NPVAC
ID VP43_NPVAC STANDARD; PRT; 363 AA.
AC P34050;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE 43 KD PROTEIN.
GN P43.
OS AUTOGRAHA CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS (ACMPNV).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; BACULOVIRIDAE;
OC NUCLEOPOLYHEDROVIRUS.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=HR3;
RX MEDLINE; 93233214.
RX CARSTENS E.B., LU A.L., CHAN H.L.B.;
RT "Sequence, transcriptional mapping, and overexpression of p47, a
RT baculovirus gene regulating late gene expression."
RL J. VIROL. 67:2513-2520(1993).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE; 94303173.
RX AYRES M.D., HOWARD S.C., KUZIO J., LOPEZ-FERBER M., POSSEE R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RT polyhedrosis virus."
RL VIROLOGY 202:586-605(1994).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L07878; G293991; -.
CC EMBL; L22858; G559108; -.

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:01 ; Search time 71.09 Seconds  
(without alignments)  
3.579 Million cell updates/sec

Title: US-09-124-280A-23  
Perfect score: 44  
Sequence: 1 KVRKSFFKV 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	93.2	481	1 LBP_HUMAN	P18428 homo sapien
2	35	79.5	481	1 LBP_MOUSE	Q61805 mus musculus
3	34	77.3	481	1 VP43_NPVAC	P34050 autographa
4	34	77.3	249	1 Y129_BORBU	O51155 borrelia bu
5	33	75.0	851	1 OPPE_MYCPN	P75551 mycoplasma
6	33	75.0	217	1 RS5_MYCLE	O33000 mycobacteri
7	33	75.0	92	1 UCRO_HUMAN	O14949 homo sapien
8	31	70.5	566	1 KAD5_SCHPO	Q09831 schizosacch
9	31	70.5	481	1 LBP_RAT	Q63313 rattus norv
10	31	70.5	713	1 PALLY_RHORB	P10248 rhodotorula
11	31	70.5	740	1 PFL_CLOPA	Q46266 clostridium
12	31	70.5	773	1 PIGR_RABIT	P01832 oryctolagus
13	31	70.5	48	1 RK32_VICFA	P15820 vicia faba
14	31	70.5	196	1 RL11_CAEFL	Q94300 caenorhabdi
15	31	70.5	130	1 RR11_ODOSI	P49499 odontella s
16	31	70.5	431	1 YAZI_SCHPO	O10223 schizosacch
17	31	70.5	166	1 YPI6_CLOPE	P18017 clostridium
18	30	68.2	445	1 BPI_RABIT	Q28739 oryctolagus
19	30	68.2	1235	1 DPOL_PYRHO	O59610 pyrococcus
20	30	68.2	1312	1 DPOL_PYRSD	O51334 pyrococcus
21	30	68.2	1702	1 DPOL_THERLI	P30317 thermococcu
22	30	68.2	1829	1 DPOL_THERST	O33845 thermococcu
23	30	68.2	1280	1 MDRI_HUMAN	P08183 homo sapien
24	30	68.2	544	1 MT57_ECOLI	P25240 escherichia
25	30	68.2	183	1 RS5_MICLU	P33105 micrococcus
26	30	68.2	328	1 YIAO_HAEIN	P44992 haemophilus
27	30	68.2	320	1 YPCP_BACSU	P39791 bacillus su
28	29	65.9	3329	1 BRC2_MOUSE	P97929 mus musculu
29	29	65.9	90	1 IFI_BORBU	O51191 borrelia bu
30	29	65.9	482	1 LBP_RABIT	P17454 oryctolagus
31	29	65.9	264	1 PCN2_CHVP1	O41056 parametium
32	29	65.9	350	1 REBB_RHINI	P55462 rhizobium s
33	29	65.9	362	1 RT04_ARATH	Q31708 arabidopsis
34	29	65.9	570	1 SVX_ARCFU	Q29346 archaeglob
35	29	65.9	473	1 TRAI_CLOPE	Q05309 clostridium
36	29	65.9	235	1 YIQ1_YEAST	P40449 saccharomyc
37	29	65.9	117	1 YIQ3_YEAST	P40448 saccharomyc
38	29	65.9	858	1 YNU1_YEAST	P40164 saccharomyc
39	28	63.6	298	1 ADT1_RAT	Q05962 rattus norv
40	28	63.6	342	1 BONZ_CERAE	O18983 cercopithec
41	28	63.6	342	1 BONZ_HUMAN	O00574 homo sapien
42	28	63.6	342	1 BONZ_MACNE	O19024 macaca neme
43	28	63.6	482	1 BPI_BOVIN	P17453 bos taurus

ALIGNMENTS

RESULT 1

LBP\_HUMAN  
ID LBP\_HUMAN STANDARD; PRT; 481 AA.  
AC P18428; Q92672; O43438;  
DT 01-NOV-1990 (REL. 16, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).  
GN LBP.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 90385281.  
RA SCHUMANN R.R., LEONG S.R., FLAGGS G.W., GRAY P.W., WRIGHT S.D.,  
RA MATHISON J.C., TOBIAS P.S., ULEVITCH R.J.;  
RT "Structure and function of lipopolysaccharide binding protein.";  
RL SCIENCE 249:1429-1431(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA HUBACEK J.A., ASLANIDIS C., SCHMITZ G.;  
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 98110577.  
RA KIRSCHING C.J., AU-YOUNG J., LAMPING N., REUTER D., PFEIL D.,  
RA SEILHAMER J.J., SCHUMANN R.R.;  
RT "Similar organization of the lipopolysaccharide-binding protein (LBP)  
and phospholipid transfer protein (PLTP) genes suggests a common gene  
family of lipid-binding proteins.";  
RL GENOMICS 46:416-425(1997).  
RN [4]  
RP SEQUENCE OF 1-41 FROM N.A.  
RA SUTTON C.L., SMITH R.I.F., CENTOLA M.B., THEOFAN G.;  
RL SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [5]  
RP 3D-STRUCTURE MODELLING.  
RX MEDLINE: 98227852.  
RA BEAVER L.J., CARROLL S.F., EISENBERG D.;  
RT "The BPI/LBP family of proteins: a structural analysis of conserved  
regions.";  
RL PROTEIN SCI. 7:906-914(1998).  
CC -!- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL  
MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS  
TO INTERACT WITH THE CD14 RECEPTOR.  
CC -!- SIMILARITY: BELONGS TO THE BPI/CTFP/LBP/PLTP FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: M35533; G186966;  
CC EMBL: X98657; E264344;  
CC EMBL: X98658; E264344; JOINED.  
CC EMBL: X98659; E264344; JOINED.  
CC EMBL: X98660; E264344; JOINED.  
CC EMBL: X98661; E264344; JOINED.  
CC EMBL: X98662; E264344; JOINED.  
CC EMBL: X98663; E264344; JOINED.  
CC EMBL: X98664; E264344; JOINED.  
CC -----

**THIS PAGE BLANK (USPTO)**

Wed Sep 8 09:43:17 1999

us-09-124-280a-23.rpr

Page 5

Search completed: September 7, 1999, 23:22:37  
Job time: 944 sec

11  
10  
9  
8  
7  
6  
5  
4  
3  
2  
1



Db 163 QLKSFYKV 171

## RESULT 12

R5VF32

ribosomal protein L32 - fava bean chloroplast

C:Species: Chloroplast Vicia faba (fava bean)

C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 05-Sep-1997

C:Accession: S08495; S08444

R:Hardenberger, F.; Weil, J.H.; Steinmetz, A.

Curr. Genet. 14, 609-615, 1988

A:Title: Organization and nucleotide sequence of the broad bean chloroplast genes trnL-U

A:Reference number: S08494; MUID:89208884

A:Accession: S08495

A:Molecule type: DNA

A:Residues: 1-48 &lt;HER&gt;

A:Cross-references: EMBL:X51471; NID:g12387; PID:g12389

C:Genetics:

A:Genome: chloroplast

C:Superfamily: rice chloroplast ribosomal protein L32

C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 70.5%; Score 31; DB 1; Length 48;

Best Local Similarity 62.5%; Pred. No. 5.1;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFK 8

I:|||||

Db 14 KIRKFWK 21

## RESULT 13

JC6010

formate C-acetyltransferase (EC 2.3.1.54) - Clostridium pasteurianum

N:Alternate names: Pfl protein; pyruvate formate-lyase

C:Species: Clostridium pasteurianum

C:Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 06-Nov-1998

C:Accession: JC6010; PC6004

R:Weidner, G.; Sawers, G.

J. Bacteriol. 178, 2440-2444, 1996

A:Title: Molecular characterization of the genes encoding pyruvate formate-lyase and its

A:Reference number: JC6010; MUID:96218720

A:Accession: JC6010

A:Molecule type: DNA

A:Residues: 1-740 &lt;WEI1&gt;

A:Cross-references: EMBL:X93463; NID:g1072360; PID:g1072361

A:Accession: PC6004

A:Molecule type: protein

A:Residues: 1-740 &lt;WEI2&gt;

C:Comment: This enzyme must be activated by the lyase-specific pyruvate formate-lyase-ac

ite and causes peptide cleavage.

C:Genetics:

A:Gene: pfl

A:Start codon: TTG

C:Complex: homodimer

C:Function:

A:Description: catalyzes the reversible conversion of acetyl-CoA and formate into pyruva

A:Pathway: anaerobic glucose metabolism

C:Superfamily: formate C-acetyltransferase 1; glycyl radical homology

C:Keywords: acyltransferase; homodimer; lyase; thiolester bond

F:681-740/Domain: glycyl radical homology &lt;GFR&gt;

F:405/Active site: Cys (cysteine thyl radical intermediate) #status predicted

F:406/Active site: Cys (S-acetylcysteine intermediate) #status predicted

F:715/Active site: Gly (stable glycyl radical) #status predicted

Query Match 70.5%; Score 31; DB 2; Length 740;

Best Local Similarity 55.6%; Pred. No. 81;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFKV 9

I:|||||

Db 461 KVENYFKV 469

## RESULT 14

S01999

phenylalanine ammonia-lyase (EC 4.3.1.5) - fungus (Rhodotorula rubra)

C:Species: Rhodotorula rubra

C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 21-Aug-1998

C:Accession: S01999; S01990

R:Filpula, D.; Vaslet, C.A.; Levy, A.; Sykes, A.; Strausberg, R.L.

Nucleic Acids Res. 16, 11381, 1988

A:Title: Nucleotide sequence of gene for phenylalanine ammonia-lyase from Rhodotorula

A:Reference number: S01999; MUID:89083582

A:Accession: S01999

A:Molecule type: DNA

A:Residues: 1-713 &lt;FIL&gt;

A:Cross-references: EMBL:X13094; NID:g3284; PID:g295942

R:Vaslet, C.A.; Strausberg, R.L.; Sykes, A.; Levy, A.; Filpula, D.

Nucleic Acids Res. 16, 11382, 1988

A:Title: cDNA and genomic cloning of yeast phenylalanine ammonia-lyase genes reveal g

A:Reference number: S01990; MUID:89083583

A:Accession: S01990

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-674 &lt;VAS&gt;

A:Cross-references: EMBL:X13095

C:Genetics:

A:Introns: 110/2; 139/2; 372/3; 428/2; 482/3

C:Superfamily: histidine ammonia-lyase

C:Keywords: ammonia-lyase; carbon-nitrogen lyase

F:218/Modified site: serine derivative (Ser) #status predicted

Query Match 70.5%; Score 31; DB 2; Length 713;

Best Local Similarity 75.0%; Pred. No. 78;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KVRKSFFK 8

I:|||||

Db 573 KVRKSIYK 580

## RESULT 15

S78255

ribosomal protein S11, chloroplast - Odontella sinensis chloroplast

C:Species: chloroplast Odontella sinensis

C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 24-Apr-1998

C:Accession: S78255

R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sine

A:Reference number: S78238

A:Accession: S78255

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-130 &lt;ROW&gt;

A:Cross-references: EMBL:267753; NID:g1185127; PID:e211821; PID:g1185145

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

C:Genetics:

A:Gene: rps11

A:Genome: chloroplast

C:Superfamily: Escherichia coli ribosomal protein S11

C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 70.5%; Score 31; DB 2; Length 130;

Best Local Similarity 75.0%; Pred. No. 14;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KVRKSFFK 8

I:|||||

Db 4 KIRKSTFK 11

S73939  
oligopeptide transport ATP-binding protein oppF - Mycoplasma pneumoniae (ATCC 29342) (SG)  
N:Alternate names: hypothetical protein G07\_0f851  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 17-Jul-1998  
C:Accession: S73939  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327; MUID:97105885  
A:Accession: S73939  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-851 <HIM>  
A:Cross-references: EMBL:AE000058; GB:U00089; NID:g1674291; PID:g1674315  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Gene: oppF  
A:Genetic code: SGC3  
C:Superfamily: ATP-binding cassette homology  
C:Keywords: P-loop  
F:31-764/Domain: ATP-binding cassette homology #status atypical <ABCL>  
F:48-55/Region: nucleotide-binding motif A (P-loop)

Query Match 75.0%; Score 33; DB 2; Length 851;  
Best Local Similarity 87.5%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFFK 8  
||| ||||  
DB 22 KVRGSFFK 29

RESULT 9  
S71288  
protoporphyrin IX magnesium chelatase (EC 4.99.1.1) chlH - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 18-Sep-1998  
C:Accession: S71288  
R:Gibson, L.; Harrison, J.; Leech, R.; Jensen, P.; Bassham, D.; Gibson, M.; Hunter, C.  
submitted to the EMBL Data Library, January 1996  
A:Description: A putative Mg chelatase subunit from Arabidopsis thaliana cv. C24: sequenced transcript.  
A:Reference number: S71288  
A:Accession: S71288  
A:Molecule type: DNA  
A:Residues: 1-1381 <GIB>  
A:Cross-references: EMBL:Z68495; NID:g1154626; PID:e218012; PID:g1154627  
A:Experimental source: cv. C24  
A:Note: probably with transit peptide  
C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase  
C:Keywords: lyase; magnesium

Query Match 72.7%; Score 32; DB 2; Length 1381;  
Best Local Similarity 77.8%; Pred. No. 93;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KVRKSFFK 9  
| |||||  
DB 39 KPAKSFFK 47

RESULT 10  
F70448  
conserved hypothetical protein aq\_1723 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 08-May-1998  
C:Accession: F70448  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy

Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: F70448  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-385 <AQF>  
A:Cross-references: GB:AE000753; NID:g2984035; PID:g2984044; GB:AE000657  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: aq\_1723

Query Match 72.7%; Score 32; DB 2; Length 385;  
Best Local Similarity 85.7%; Pred. No. 26;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRKSFFK 8  
||| ||||  
DB 339 VRKAFFK 345

RESULT 11  
QRRBG  
secretory component precursor - rabbit  
N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor  
N:Contains: free secretory component; transmembrane secretory component  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 15-Nov-1984 #sequence\_revision 15-Nov-1984 #text\_change 25-Oct-1996  
C:Accession: A02111; A28077  
R:Mostov, K.E.; Friedlander, M.; Blobel, G.  
Nature 308, 37-43, 1984  
A:Title: The receptor for trans epithelial transport of IgA and IgM contains multiple  
A:Reference number: A02111; MUID:84142246  
A:Accession: A02111  
A:Molecule type: mRNA  
A:Residues: 1-773 <MOS>  
A:Note: the authors translated the codon ACC for residue 54 as Asn  
R:Frutiger, S.; Hughes, G.J.; Hanly, W.C.; Jaton, J.C.  
J. Biol. Chem. 263, 8120-8125, 1988  
A:Title: Rabbit secretory components of different allotypes vary in their carbohydrate  
A:Reference number: A28077; MUID:88228032  
A:Accession: A28077  
A:Molecule type: protein  
A:Residues: 87-114;410-424 <FRU>  
C:Comment: This receptor binds polymeric IgA and IgM at the basolateral surface of ep  
process, cleavage occurs to separate the extracellular portion, also known as the sec  
C:Comment: The five domains exhibit homology with immunoglobulin V regions. The simil  
C:Comment: Alternative splicing in the extracellular domain leads to high or low mole  
C:Keywords: secretory component; immunoglobulin homology  
C:Superfamily: secretory component; immunoglobulin homology  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-773/Product: transmembrane secretory component #status predicted <MATM>  
F:19-575/Product: free secretory component #status predicted <MATP>  
F:30-647/Domain: extracellular #status predicted <EXT>  
F:39-117/Domain: immunoglobulin homology <IGV1>  
F:148-227/Domain: immunoglobulin homology <IGV2>  
F:253-326/Domain: immunoglobulin homology <IGV3>  
F:362-440/Domain: immunoglobulin homology <IGV4>  
F:471-540/Domain: immunoglobulin homology <IGV5>  
F:648-670/Domain: transmembrane #status predicted <TMN>  
F:671-773/Domain: intracellular #status predicted <INT>  
F:46-115,155-225,260-324,369-438,478-538/Disulfide bonds: #status predicted  
F:108/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:418/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 70.5%; Score 31; DB 1; Length 773;  
Best Local Similarity 55.6%; Pred. No. 84;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFK 9  
::: ||||: ||

variable outer membrane protein - Borrelia hermslii

C:Species: Borrelia hermslii  
C:Date: 09-Mar-1996 #sequence\_revision 09-Mar-1996 #text\_change 09-Mar-1996  
C:Accession: I40309

R:Restrepo, B. I.; Barbour, A. G.

Cell 78, 867-876, 1994

A:Title: Antigen diversity in the bacterium B. hermslii through 'somatic' mutations in re

A:Reference number: A54824; MUID:94373822

A:Accession: I40309

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <RES>

A:Cross-references: GB:L33898; NID:g558328; PID:g558329

C:Genetics:

A:Gene: vmp6

Query Match 79.5%; Score 35; DB 2; Length 215;  
Best Local Similarity 77.8%; Pred. No. 3.4;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9

|||||

DB 172 KVRKSLFKL 180

RESULT 4

A45684

orf2\_3' of p47 - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, ACMNPV

C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994

C:Accession: A45684

R:Carstens, E. B.; Lu, A. L.; Chan, H. L.

J. Virol. 67, 2513-2520, 1993

A:Title: Sequence, transcriptional mapping, and overexpression of p47, a baculovirus gen

A:Reference number: A45684; MUID:93233214

A:Accession: A45684

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-363 <CAR>

A>Note: sequence extracted from NCBI backbone (NCBIN:129587, NCBI:129588)

Query Match 77.3%; Score 34; DB 2; Length 363;

Best Local Similarity 75.0%; Pred. No. 9.3;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFK 8

|||||

DB 313 KVRKSFFK 320

RESULT 5

A70116

Conserved hypothetical protein BB0129 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998

C:Accession: A70116

R:Fraser, C. M.; Castjens, S.; Huang, W. M.; Sutton, G. G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kerlavage, A. R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M. D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H. O.; Venter, J. C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943

A:Accession: A70116

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-249 <KLE>

A:Cross-references: GB:AE001124; GB:AE000783; NID:g2688003; PID:g2688006; TIGR:BB0129

A:Experimental source: strain B31

Query Match 77.3%; Score 34; DB 2; Length 249;

Best Local Similarity 66.7%; Pred. No. 6.3;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9

|||||

DB 163 KVRKEFFKL 171

RESULT 6

D70644

Probable rpsE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 21-Nov-1998

C:Accession: D70644

R:Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: D70644

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-220 <COL>

A:Cross-references: GB:Z84395; GB:AL123456; NID:g3261698; PID:e293131; PID:g1806189

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: rpsE

C:Superfamily: Escherichia coli ribosomal protein S5

Query Match 75.0%; Score 33; DB 2; Length 220;

Best Local Similarity 66.7%; Pred. No. 9.1;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9

|||||

DB 97 EARKSFFRV 105

RESULT 7

B64657

Conserved hypothetical secreted protein HP1098 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence\_revision 18-Sep-1997 #text\_change 10-Oct-1997

C:Accession: B64657

R:Tomb, J. F.; White, O.; Kerlavage, A. R.; Clayton, R. A.; Sutton, G. G.; Fleischmann, R

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H. G.; Glodek, A.; McKe

son, J. D.; Kelley, J. M.; Cotton, M. D.; Weidman, J. M.; Fujii, C.; Bowman, C.; Watthey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W. S.; Borodovsky, M.; Karpk, P. D.; Smith, H. O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: B64657

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-290 <TOM>

A:Cross-references: GB:AE000511; NID:g2314242; PID:g2314247; TIGR:HP1098

C:Superfamily: conserved hypothetical secreted protein HP1098

Query Match 75.0%; Score 33; DB 2; Length 290;

Best Local Similarity 75.0%; Pred. No. 12;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRKSFFKV 9

|||||

DB 5 VRKSFFRV 12

RESULT 8

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:35 ; Search time 105.14 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-23  
Perfect score: 44  
Sequence: 1 KVRKSFVKV 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	93.2	481	A54136	lipopolysaccharide
2	41	93.2	477	A35843	lipopolysaccharide
3	35	79.5	215	I40309	variable outer mem
4	34	77.3	363	A45684	orf2 3' of p47 - A
5	34	77.3	249	A70116	conserved hypothet
6	33	75.0	220	D70644	probable rpsE prot
7	33	75.0	230	B64657	conserved hypothet
8	33	75.0	851	S73939	oligopeptide trans
9	32	72.7	1381	S71288	protoporphyrin IX
10	32	72.7	385	F70448	conserved hypothet
11	31	70.5	773	ORR8G	secretory componen
12	31	70.5	48	R5VP32	ribosomal protein
13	31	70.5	740	JC6010	formate C-acetyltr
14	31	70.5	713	S01999	phenylalanine ammo
15	31	70.5	130	S78255	ribosomal protein
16	31	70.5	481	I56246	lipopolysaccharide
17	31	70.5	566	S62482	hypothetical prote
18	31	70.5	524	B70479	2-isopropylmalate
19	31	70.5	362	JQ1602	replication protei
20	31	70.5	186	JT0358	hypothetical prote
21	31	70.5	1021	T00361	hypothetical prote
22	31	70.5	678	F70451	hypothetical prote
23	30	68.2	1280	DVH01	multidrug resistan
24	30	68.2	1702	S42459	DNA-directed DNA p
25	30	68.2	183	S29888	Ribosomal protein
26	30	68.2	1235	C71210	probable DNA-direc
27	30	68.2	1312	S68593	DNA-directed DNA p
28	30	68.2	232	F70344	endonuclease III -
29	30	68.2	328	I64164	hypothetical prote
30	30	68.2	437	F70114	probable zinc prot
31	30	68.2	333	C64664	transferase, pepti
32	30	68.2	353	G71852	udp-n-acetylglucos
33	30	68.2	544	S26425	site-specific DNA-
34	30	68.2	320	H69939	hypothetical prote
35	30	68.2	98	A69262	hypothetical prote
36	30	68.2	99	T01915	translation initia
37	29	65.9	435	C70456	homoserine dehydro
38	29	65.9	524	A45472	protein kinase (EC
39	29	65.9	570	D69364	glycyl-tRNA synthe

ALIGNMENTS

RESULT 1

A54136  
lipopolysaccharide-binding protein - human  
N:Alternate names: LBP; LPS-binding protein  
C:Species: Homo sapiens (man)  
C:Date: 11-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 15-Jun-1996  
C:Accession: A54136  
R:Wilide, C.G.; Seilhamer, J.J.; McGrogan, M.; Ashton, N.; Snable, J.L.; Lane, J.C.; L  
J. Biol. Chem. 269, 17411-17416, 1994  
A:Title: Bactericidal/permeability-increasing protein and lipopolysaccharide (LPS)-bi  
A:Reference number: A54136; MUID:94292492  
A:Accession: A54136  
A:Status: Preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-481 <SWIL>  
A:Experimental source: liver  
A>Note: sequence extracted from NCBI backbone (NCBIP:149401)  
C:Superfamily: lipopolysaccharide-binding protein

Query Match 93.2%; Score 41; DB 2; Length 481;  
Best Local Similarity 88.9%; Pred. No. 0.42;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFVKV 9  
|||  
Db 117 KVRKSFVKL 125

RESULT 2

A35843  
lipopolysaccharide-binding protein - human  
C:Species: Homo sapiens (man)  
C:Date: 05-Oct-1990 #sequence\_revision 05-Oct-1990 #text\_change 24-Sep-1998  
C:Accession: A35843  
R:Schumann, R.R.; Leong, S.R.; Flagg, G.W.; Gray, P.W.; Wright, S.D.; Mathison, J.C.  
Science 249, 1429-1431, 1990  
A:Title: Structure and function of lipopolysaccharide binding protein.  
A:Reference number: A35843; MUID:90385281  
A:Accession: A35843  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-477 <SCH>  
A:Cross-references: GB:M35533; NID:gl86965; PID:gl86966  
C:Genetics:  
A:Gene: GDB:LBP  
A:Cross-references: GDB:131571; OMIM:151990  
A:Map position: 20q11.23-20q12  
C:Superfamily: lipopolysaccharide-binding protein

Query Match 93.2%; Score 41; DB 2; Length 477;  
Best Local Similarity 88.9%; Pred. No. 0.42;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFVKV 9  
|||  
Db 117 KVRKSFVKL 125

RESULT 3

I40309

**THIS PAGE BLANK (USPTO)**

us-09-124-280a-23.ra1

Wed Sep 8 09:43:16 1999

```

; TITLE OF INVENTION: Lipopolysaccharide Binding Protein Derivatives
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,660A
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/32137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "rlbp25"
; US-08-261-660A-45

```

```

Query Match          93.28; Score 41; DB 1; Length 197;
Best Local Similarity 88.9%; Pred. No. 0.82;
Matches      8; Conservative 1; Mismatches 0; Indels 0; Caps 0;

```

```

Qy      1 KVRKSFFKV 9
        |||||
Db      92 KVRKSFFKL 100

```

Search completed: September 7, 1999, 23:07:45  
Job time: 1743 sec

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/32137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "rlbp25"
;
US-08-261-660A-2

Query Match          93.2%; Score 41; DB 1; Length 197;
Best Local Similarity 88.9%; Pred. No. 0.82;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9
   |||||
Db 92 KVRKSEFKL 100

RESULT 13
US-08-261-660A-4
; Sequence 4, Application US/08261660A
; Patent No. 5731415
; GENERAL INFORMATION:
; APPLICANT: Gazzano-Santoro, Helene
; APPLICANT: Theofan, Georgia
; APPLICANT: Trown, Patrick W.
; TITLE OF INVENTION: Lipopolysaccharide Binding Protein Derivatives
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/32137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "rlbp"
```

```
US-08-261-660A-4

Query Match          93.2%; Score 41; DB 1; Length 481;
Best Local Similarity 88.9%; Pred. No. 1.7;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9
   |||||
Db 117 KVRKSEFKL 125

RESULT 14
US-08-261-660A-29
; Sequence 29, Application US/08261660A
; Patent No. 5731415
; GENERAL INFORMATION:
; APPLICANT: Gazzano-Santoro, Helene
; APPLICANT: Theofan, Georgia
; APPLICANT: Trown, Patrick W.
; TITLE OF INVENTION: Lipopolysaccharide Binding Protein Derivatives
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/32137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-261-660A-29

Query Match          93.2%; Score 41; DB 1; Length 13;
Best Local Similarity 88.9%; Pred. No. 0.088;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9
   |||||
Db 5 KVRKSEFKL 13

RESULT 15
US-08-261-660A-45
; Sequence 45, Application US/08261660A
; Patent No. 5731415
; GENERAL INFORMATION:
; APPLICANT: Gazzano-Santoro, Helene
; APPLICANT: Theofan, Georgia
; APPLICANT: Trown, Patrick W.
```

us-09-124-280a-23.ra1

Wed Sep 8 09:43:16 1999

```

; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "ILBP"
; US-08-306-473A-98

Query Match 93.2%; Score 41; DB 1; Length 481;
Best Local Similarity 88.9%; Pred. No. 1.7;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9
DB 117 KVRKSEFKL 125

RESULT 10
US-08-240-328-2
; Sequence 2, Application US/08240328
; Patent No. 5698680
; GENERAL INFORMATION:
; APPLICANT: Ulevitch, Richard
; TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING
; TITLE OF INVENTION: ENDOXIC ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240.328
; FILING DATE: 14-MAY-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: LISA A. TUMARKIN, PH.D.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD-3575
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-240-328-2

Query Match 93.2%; Score 41; DB 1; Length 477;
Best Local Similarity 88.9%; Pred. No. 1.7;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9
DB 117 KVRKSEFKL 125

RESULT 11
US-08-205-719-3
; Sequence 3, Application US/08205719
; Patent No. 5705398
; GENERAL INFORMATION:
; APPLICANT: Mintz, D. N.
; APPLICANT: Tobias, P. S.
; APPLICANT: Ulevitch, R. J.
; TITLE OF INVENTION: SYSTEM AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: THERAPEUTIC INHIBITORS OF LPS-MEDIATED SEPSIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5705398th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/205,719
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SRI201P; TSRI324.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-205-719-3

Query Match 93.2%; Score 41; DB 1; Length 452;
Best Local Similarity 88.9%; Pred. No. 1.6;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFKV 9
DB 92 KVRKSEFKL 100

RESULT 12
US-08-261-660A-2
; Sequence 2, Application US/08261660A
; Patent No. 5731415
; GENERAL INFORMATION:
; APPLICANT: Gazzano-Santoro, Helene
; APPLICANT: Theofan, Georgia
; APPLICANT: Town, Patrick W.
; TITLE OF INVENTION: Lipopolysaccharide Binding Protein Derivatives
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,660A
; FILING DATE: 17-JUN-1994
```



;  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/372,105  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,540  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,762  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,222  
; FILING DATE: 14-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27129/32415  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 481 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "rLBP"  
US-08-372-105-98

Query Match 93.2%; Score 41; DB 1; Length 481;  
Best Local Similarity 88.9%; Pred. No. 1.7;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFVKV 9  
Db 117 KVRKSFVKL 125

RESULT 8  
US-08-097-830E-23  
; Sequence 23, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E

;  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-23

Query Match 93.2%; Score 41; DB 1; Length 9;  
Best Local Similarity 88.9%; Pred. No. 7.9e+04;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFVKV 9  
Db 1 KVRKSFVKL 9

RESULT 9  
US-08-306-473A-98  
; Sequence 98, Application US/08306473A  
; Patent No. 5652332  
; GENERAL INFORMATION:  
; APPLICANT: Little, Roger G.  
; TITLE OF INVENTION: Biologically Active Peptides from  
; TITLE OF INVENTION: Functional Domains of Bactericidal/  
; TITLE OF INVENTION: Permeability-Increasing Protein and  
; TITLE OF INVENTION: Uses Thereof  
; NUMBER OF SEQUENCES: 226  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Allegretti & Witcoff, Ltd.  
; STREET: Suite 3000, 10 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306,473A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,762  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,222  
; FILING DATE: 14-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McDonnell, John J.  
; REGISTRATION NUMBER: 26,949  
; REFERENCE/DOCKET NUMBER: 93,1133-  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 481 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

QY 1 KVRKSFFKV 9  
Db 92 KVRKSFFKL 100

## RESULT 5

US-08-311-611A-98  
; Sequence 98, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 481 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "rLBP"  
US-08-311-611A-98

Query Match 93.2%; Score 41; DB 1; Length 481;  
Best Local Similarity 88.9%; Pred. No. 1.7;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9  
Db 117 KVRKSFFKL 125

## RESULT 6

US-08-372-783-98  
; Sequence 98, Application US/08372783  
; Patent No. 5578572  
; GENERAL INFORMATION:

; APPLICANT: Horwitz, Arnold H.  
; APPLICANT: Lambert, Lewis H.  
; APPLICANT: Little, Roger G.  
; TITLE OF INVENTION: Anti-Gram-Positive Bacterial Methods and  
; TITLE OF INVENTION: Materials  
; NUMBER OF SEQUENCES: 237  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/372,783  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,540  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,762  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,222  
; FILING DATE: 14-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27129/32415  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 481 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "rLBP"  
US-08-372-783-98

Query Match 93.2%; Score 41; DB 1; Length 481;  
Best Local Similarity 88.9%; Pred. No. 1.7;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9  
Db 117 KVRKSFFKL 125

## RESULT 7

US-08-372-105-98  
; Sequence 98, Application US/08372105  
; Patent No. 5627153  
; GENERAL INFORMATION:  
; APPLICANT: Little, Roger G.  
; APPLICANT: Lim, Edward  
; APPLICANT: Lambert, Lewis H.  
; APPLICANT: Scannon, Patrick J.  
; TITLE OF INVENTION: Anti-Fungal Materials and Methods  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Spensley Horn Jubas & Lubitz  
;; STREET: 4225 Executive Square, Suite 1400  
;; CITY: La Jolla  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92037  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/847,562  
;; FILING DATE: 19920302  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Wetherell, Jr. Ph. D. John R.  
;; REGISTRATION NUMBER: 31,678  
;; REFERENCE/DOCKET NUMBER: PD-1859  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 455-5100  
;; TELEFAX: (619) 455-5110  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 477 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-07-847-562-2

Query Match 93.2%; Score 41; DB 1; Length 477;  
Best Local Similarity 88.9%; Pred. No. 1.7;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9

Db 117 KVRKSFFKL 125

### RESULT 3

US-08-186-811-2  
; Sequence 2, Application US/08186811  
; Patent No. 5484705  
; GENERAL INFORMATION:  
; APPLICANT: White, Mark Leslie  
; APPLICANT: Carroll, Stephen Fitzhugh  
; TITLE OF INVENTION: METHOD FOR QUANTIFYING LBP IN BODY FLUIDS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,811  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 27129/31843  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300

;; TELEFAX: 312/474-0448  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 481 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "rlbp"  
;; US-08-186-811-2

Query Match 93.2%; Score 41; DB 1; Length 481;  
Best Local Similarity 88.9%; Pred. No. 1.7;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9

Db 117 KVRKSFFKL 125

### RESULT 4

US-08-186-811-4  
; Sequence 4, Application US/08186811  
; Patent No. 5484705  
; GENERAL INFORMATION:  
; APPLICANT: White, Mark Leslie  
; APPLICANT: Carroll, Stephen Fitzhugh  
; APPLICANT: Ma, Jeremy Kam-Kuen  
; TITLE OF INVENTION: METHOD FOR QUANTIFYING LBP IN BODY FLUIDS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,811  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 27129/31843  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 197 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "rlbp25"  
; US-08-186-811-4

Query Match 93.2%; Score 41; DB 1; Length 197;  
Best Local Similarity 88.9%; Pred. No. 0.82;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:44 ; Search time 82.37 Seconds.  
(without alignments)  
1.078 Million cell updates/sec

Title: US-09-124-280A-23  
Perfect score: 44  
Sequence: 1 KVRKSEFFKV 9

Scoring table: BLOSUM62

Searched: 105577 seqs, 9869381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2.6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2.6/ptodata/2/1aa/PCTUS9\_COMB.pep:\*  
4: /cgn2.6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	9	2	US-08-456-112B-23
2	41	93.2	477	1	US-07-847-562-2
3	41	93.2	481	1	US-08-186-811-2
4	41	93.2	197	1	US-08-186-811-4
5	41	93.2	481	1	US-08-311-611A-98
6	41	93.2	481	1	US-08-372-783-98
7	41	93.2	481	1	US-08-372-108-98
8	41	93.2	9	1	US-08-097-830E-23
9	41	93.2	481	1	US-08-306-473A-98
10	41	93.2	477	1	US-08-240-328-2
11	41	93.2	452	1	US-08-205-719-3
12	41	93.2	197	1	US-08-261-660A-2
13	41	93.2	481	1	US-08-261-660A-4
14	41	93.2	13	1	US-08-261-660A-29
15	41	93.2	197	1	US-08-261-660A-45
16	41	93.2	198	1	US-08-261-660A-50
17	41	93.2	198	1	US-08-261-660A-52
18	41	93.2	199	1	US-08-261-660A-54
19	41	93.2	198	1	US-08-261-660A-56
20	41	93.2	481	1	US-08-209-762-98
21	41	93.2	481	2	US-08-473-344-98
22	41	93.2	197	2	US-08-274-303-4
23	41	93.2	481	2	US-08-274-303-6
24	41	93.2	454	2	US-07-915-720D-17
25	41	93.2	481	2	US-08-377-391A-2
26	41	93.2	197	2	US-08-377-391A-4
27	41	93.2	197	2	US-08-215-089-2
28	41	93.2	481	2	US-08-215-089-9
29	41	93.2	481	2	US-08-485-445A-98
30	41	93.2	477	2	US-08-990-849-2
31	41	93.2	481	3	PCT-US94-02465-98
32	41	93.2	197	3	PCT-US94-06931-2
33	41	93.2	481	3	PCT-US94-06931-4
34	41	93.2	13	3	PCT-US94-06931-29
35	41	93.2	197	3	PCT-US94-07834-4
36	41	93.2	481	3	PCT-US94-07834-6
37	41	93.2	481	3	PCT-US95-00498-98
38	41	93.2	481	3	PCT-US95-00656-98
39	41	93.2	197	3	PCT-US95-03384-2

40 41 93.2 481 3 PCT-US95-03384-9  
41 40 90.9 27 1 US-08-372-783-236  
42 40 90.9 35 1 US-08-261-660A-21  
43 40 90.9 14 1 US-08-261-660A-38  
44 40 90.9 27 1 US-08-261-660A-43  
45 40 90.9 35 3 PCT-US94-06931-21

ALIGNMENTS

RESULT 1  
US-08-456-112B-23  
; Sequence 23, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-23

Query Match 100.0%; Score 44; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSEFFKV 9  
| | | | | | | | |  
DB 1 KVRKSEFFKV 9

RESULT 2  
US-07-847-562-2  
; Sequence 2, Application US/07847562  
; Patent No. 5310879  
; GENERAL INFORMATION:  
; APPLICANT: Ulevitch, Richard  
; APPLICANT: Tobias, Peter  
; TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC  
; TITLE OF INVENTION: ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND  
; TITLE OF INVENTION: POLYPEPTIDES  
; NUMBER OF SEQUENCES: 2

Best Local Similarity 88.9%; Pred. No. 0.5;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSEFKV 9  
          |||||||  
Db 92 KVRKSEFKL 100

Search completed: September 7, 1999, 22:49:56  
Job time: 7877 sec

AC R68930;  
 DT 11-SEP-1995 (first entry)  
 DE LBP(1-146)/BPI(148-161)/LPB(160-197) hybrid protein.  
 KW Lipopolysaccharide binding protein; hybrid protein;  
 KW LBP(1-146)/BPI(148-161)/LPB(160-197); LPS; treatment;  
 KW gram-negative bacterial infections;  
 KW bactericidal/permeability-increasing protein.  
 OS Homo sapiens.  
 PN WO9500641-A.  
 PD 05-JAN-1995.  
 PF 17-JUN-1994; U06931.  
 PR 17-JUN-1993; US-079510.  
 PA (XOMA ) XOMA CORP.  
 PI Gazzano-santoro H, Theofan G, Trown PW;  
 DR WPI; 95-052078/07.  
 PT Lipo.polysaccharide binding protein deriv. and hybrid protein  
 PT binds to lipo.polysaccharide - lacks CD14-mediated  
 PT immuno:stimulatory properties, used to treat Gram-negative  
 PT bacterial infections and associated conditions  
 PS Claim 9; Page 87; 114pp; English.  
 CC R68924-R68934 are lipopolysaccharide (LPS) binding protein (LBP)/  
 CC bactericidal/permeability increasing protein (BPI) hybrid proteins.  
 CC These hybrid proteins lack CD14-mediated immunostimulatory  
 CC properties, and are therefore used in the treatment of gram-negative  
 CC bacterial infections and associated conditions.  
 CC Sequence 198 AA;  
 SQ

Query Match 93.2%; Score 41; DB 1; Length 198;  
 Best Local Similarity 88.9%; Pred. No. 0.51;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVRKSFVKV 9  
 |||||  
 DB 92 KVRKSFVKL 100

RESULT 13  
 R68932  
 ID R68932 standard; Protein; 199 AA.  
 AC R68932;  
 DT 11-SEP-1995 (first entry)  
 DE BPI(1-85)/lipopolysaccharide BPI(86-99)/BPI(100-199) hybrid protein.  
 KW Lipopolysaccharide binding protein; BPI(1-85)/LPB(86-99)/BPI(100-199);  
 KW LPS; treatment; gram-negative bacterial infections;  
 KW hybrid protein; bactericidal/permeability-increasing protein.  
 OS Homo sapiens.  
 PN WO9500641-A.  
 PD 05-JAN-1995.  
 PF 17-JUN-1994; U06931.  
 PR 17-JUN-1993; US-079510.  
 PA (XOMA ) XOMA CORP.  
 PI Gazzano-santoro H, Theofan G, Trown PW;  
 DR WPI; 95-052078/07.  
 PT Lipo.polysaccharide binding protein deriv. and hybrid protein  
 PT binds to lipo.polysaccharide - lacks CD14-mediated  
 PT immuno:stimulatory properties, used to treat Gram-negative  
 PT bacterial infections and associated conditions  
 PS Claim 9; Page 87; 114pp; English.  
 CC R68924-R68934 are lipopolysaccharide (LPS) binding protein (LBP)/  
 CC bactericidal/permeability increasing protein (BPI) hybrid proteins.  
 CC These hybrid proteins lack CD14-mediated immunostimulatory  
 CC properties, and are therefore used in the treatment of gram-negative  
 CC bacterial infections and associated conditions.  
 CC Sequence 199 AA;  
 SQ

Query Match 93.2%; Score 41; DB 1; Length 199;  
 Best Local Similarity 88.9%; Pred. No. 0.51;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVRKSFVKV 9  
 |||||

DB 92 KVRKSFVKM 100

RESULT 14

R68934  
 ID R68934 standard; Protein; 196 AA.  
 AC R68934;  
 DT 11-SEP-1995 (first entry)  
 DE BPI(1-85)/LBP(86-99)/BPI(100-147)/LPB(147-159)/BPI(162-199).  
 KW Lipopolysaccharide binding protein; hybrid protein;  
 KW BPI(1-85)/LBP(86-99)/BPI(100-147)/LPB(147-159)/BPI(162-199);  
 KW LPS; treatment; gram-negative bacterial infections;  
 KW bactericidal/permeability-increasing protein.  
 OS Homo sapiens.  
 PN WO9500641-A.  
 PD 05-JAN-1995.  
 PF 17-JUN-1994; U06931.  
 PR 17-JUN-1993; US-079510.  
 PA (XOMA ) XOMA CORP.  
 PI Gazzano-santoro H, Theofan G, Trown PW;  
 DR WPI; 95-052078/07.  
 PT Lipo.polysaccharide binding protein deriv. and hybrid protein  
 PT binds to lipo.polysaccharide - lacks CD14-mediated  
 PT immuno:stimulatory properties, used to treat Gram-negative  
 PT bacterial infections and associated conditions  
 PS Claim 9; Page 87; 114pp; English.  
 CC R68924-R68934 are lipopolysaccharide (LPS) binding protein (LBP)/  
 CC bactericidal/permeability increasing protein (BPI) hybrid proteins.  
 CC These hybrid proteins lack CD14-mediated immunostimulatory  
 CC properties, and are therefore used in the treatment of gram-negative  
 CC bacterial infections and associated conditions.  
 CC Sequence 196 AA;  
 SQ

Query Match 93.2%; Score 41; DB 1; Length 196;  
 Best Local Similarity 88.9%; Pred. No. 0.5;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVRKSFVKV 9  
 |||||  
 DB 92 KVRKSFVKM 100

RESULT 15  
 R68915  
 ID R68915 standard; Protein; 197 AA.  
 AC R68915;  
 DT 08-SEP-1995 (first entry)  
 DE Lipopolysaccharide binding protein 1-197 (rLPB Cys131).  
 KW Lipopolysaccharide binding protein; rLPB Cys 131; LPS;  
 KW gram-negative bacterial infections; treatment.  
 OS Homo sapiens.  
 PN WO9500641-A.  
 PD 05-JAN-1995.  
 PF 17-JUN-1994; U06931.  
 PR 17-JUN-1993; US-079510.  
 PA (XOMA ) XOMA CORP.  
 PI Gazzano-santoro H, Theofan G, Trown PW;  
 DR WPI; 95-052078/07.  
 PT Lipo.polysaccharide binding protein deriv. and hybrid protein  
 PT binds to lipo.polysaccharide - lacks CD14-mediated  
 PT immuno:stimulatory properties, used to treat Gram-negative  
 PT bacterial infections and associated conditions  
 PS Claim 5; Pages 61-62; 114pp; English.  
 CC R68915 is the recombinant lipopolysaccharide (LPS) binding protein  
 CC (1-197), with a Cys for Ala 131 substitution (rLPB Cys 131). The  
 CC protein lacks CD14-mediated immunostimulatory properties, and can  
 CC therefore be used to treat gram-negative bacterial infections and  
 CC associated conditions.  
 CC Sequence 197 AA;  
 SQ

Query Match 93.2%; Score 41; DB 1; Length 197;

CC arthritis, and for treating gram-negative bacterial infection. The  
 CC peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipopolysaccharide in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipopolysaccharide or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 481 AA;

Query Match 93.2%; Score 41; DB 1; Length 481;

Best Local Similarity 88.9%; Pred. No. 1.1;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFKV 9

Db 117 KVRKSFFKL 125

RESULT 9

ID R67999 standard; Protein; 197 AA.

AC R67999;

DE 15-AUG-1995 (first entry)

DE LBP25 N-terminal sequence.

KW Bactericidal/permeability-increasing protein; BPI; rBPI; rBPI50;

KW lipopolysaccharide binding protein; LBP; rLBP25; rLBP;

KW Gram-negative bacterium; infection; antiseptic.

OS Homo sapiens.

PN W09502414-A.

PD 26-JAN-1995.

PF 13-JUL-1994; U07834.

PR 14-JUL-1993; US-093201.

PR 11-JUL-1994; US-274303.

PA (XOMA ) XOMA CORP.

PI Horwitz A;

DR WPI: 95-067161/09.

DR N-PSDB; 081445.

PT Method for treating gram negative bacterial infection - comprises

PT administering lipopolysaccharide binding protein (LBP) prod. and

PT bactericidal/permeability-increasing (BPI) protein prod.

PS Disclosure: Page 40-41; 76pp; English.

CC Gram-negative bacterial infections are treated by co-administration

CC of BPI protein and LBP. A preferred LBP is the 197-amino acid

CC N-terminal fragment (given in R67999) of recombinant LBP (rLBP,

CC R67800).

SQ Sequence 197 AA;

Query Match

Best Local Similarity 93.2%; Score 41; DB 1; Length 197;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFKV 9

Db 92 KVRKSFFKL 100

RESULT 10

ID R68000 standard; Protein; 481 AA.

AC R68000;

DE 15-AUG-1995 (first entry)

DE rLBP.

KW Bactericidal/permeability-increasing protein; BPI; rBPI; rBPI50;

KW lipopolysaccharide binding protein; LBP; rLBP25; rLBP;

KW Gram-negative bacterium; infection; antiseptic.

OS Homo sapiens.

FH Key Location/Qualifiers

FT peptide 1. .25  
 FT W09502414-A.  
 PN /label= sig\_peptide  
 PD 26-JAN-1995.  
 PF 13-JUL-1994; U07834.  
 PR 14-JUL-1993; US-093201.  
 PR 11-JUL-1994; US-274303.  
 PA (XOMA ) XOMA CORP.  
 PI Horwitz A;  
 DR WPI: 95-067161/09.  
 DR N-PSDB; 081447.

PT Method for treating gram negative bacterial infection - comprises  
 PT administering lipopolysaccharide binding protein (LBP) prod. and  
 PT bactericidal/permeability-increasing (BPI) protein prod.  
 PS Disclosure: Page 43-44; 76pp; English.  
 CC Gram-negative bacterial infections are treated by co-administration  
 CC of BPI protein and LBP. A preferred LBP is the 197-amino acid  
 CC N-terminal fragment (given in R67999) of recombinant LBP (rLBP,  
 CC R67800).

SQ Sequence 481 AA;

Query Match 93.2%; Score 41; DB 1; Length 481;

Best Local Similarity 88.9%; Pred. No. 1.1;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFKV 9

Db 117 KVRKSFFKL 125

RESULT 11

R68928

ID R68928 standard; Protein; 189 AA.

AC R68928;

DT 11-SEP-1995 (first entry)

DE BPI(1-25)/lipopolysaccharide BP(26-135)/BPI(137-199) hybrid protein.

KW Lipopolysaccharide binding protein; BPI(1-25)/LBP(26-135)/BPI(137-199);

KW LPS; treatment; gram-negative bacterial infections;

KW hybrid protein; bactericidal/permeability-increasing protein.

OS Homo sapiens.

PN W09500641-A.

PD 05-JAN-1995.

PF 17-JUN-1994; U06931.

PR 17-JUN-1993; US-079510.

PA (XOMA ) XOMA CORP.

PI Gazzano-Santoro H, Theofan G, Trown PW;

DR WPI: 95-052078/07.

PT Lipo:polysaccharide binding protein deriv. and hybrid protein

PT binds to lipo:polysaccharide - lacks CD14-mediated

PT immuno:stimulatory properties, used to treat Gram-negative

PT bacterial infections and associated conditions

PS Claim 9; Page 87; 114pp; English.

CC R68924-R68934 are lipopolysaccharide (LPS) binding protein (LBP)/

CC bactericidal/permeability increasing protein (BPI) hybrid proteins.

CC These hybrid proteins lack CD14-mediated immunostimulatory

CC properties, and are therefore used in the treatment of gram-negative

CC bacterial infections and associated conditions.

SQ Sequence 189 AA;

Query Match

Best Local Similarity 93.2%; Score 41; DB 1; Length 189;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFKV 9

Db 93 KVRKSFFKL 101

RESULT 12

R68930

ID R68930 standard; Protein; 198 AA.

CC from liver biopsies of trauma patients. The library was transferred  
 CC to nitrocellulose and screened twice with a probe based on the  
 CC N-terminal sequence of rabbit LBP. The coding sequence obtained from  
 CC positive clones encodes a glycoprotein which binds to a lipopoly-  
 CC saccharide secreted by a gram negative bacterium. The glycoprotein  
 CC retards in vitro binding of LPS to high density lipoprotein present  
 CC in normal serum.  
 SQ Sequence 477 AA;

Query Match 93.2%; Score 41; DB 1; Length 477;

Best Local Similarity 88.9%; Pred. No. 1.1;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9

|||||||:

DB 117 KVRKSFFKL 125

RESULT 6

ID R43887 standard; Protein; 197 AA.

AC R43887;

DT 01-JUN-1994 (first entry)

DE Human lipopolysaccharide binding protein.

KW Lipopolysaccharide binding protein; endotoxin binding protein.

OS Homo sapiens.

PN W09323540-A.

PD 25-NOV-1993.

PF 19-MAY-1993; U04752.

PR 19-MAY-1992; US-885501.

PA (XOMA ) XOMA CORP.

PI Grinna LS;

DR WPI; 93-386572/48.

DR P-PSDB; 052269.

PT Recovering endotoxin binding protein from transformed cell

PT cultures - by retention on cation exchange material and elution,

PT giving simpler recovery and higher yields, for treating bacterial

PT infections

PS Disclosure; Page 30-31; 39pp; English.

CC Human lipopolysaccharide binding protein (an endotoxin binding

CC protein) is useful for treatment of bacterial infections.

SQ Sequence 197 AA;

Query Match 93.2%; Score 41; DB 1; Length 197;

Best Local Similarity 88.9%; Pred. No. 0.5;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9

|||||||:

DB 92 KVRKSFFKL 100

RESULT 7

ID R53459 standard; Protein; 477 AA.

AC R53459;

DT 08-DEC-1994 (first entry)

DE Human lipopolysaccharide binding protein.

KW hLBP; human lipopolysaccharide binding protein; detection;

KW gram negative bacterial infection; ss.

OS Homo sapiens.

PN Key

FT peptide 1. .25 Location/Qualifiers

FT /label= signal peptide

FT protein 26. .477

FT /label= lipopolysaccharide binding protein

FT /note= "mature protein"

FT US5310879-A.

PN 10-MAY-1994.

PF 30-APR-1985; 728833.

PR 30-APR-1985; US-728833.

PR 30-DEC-1986; US-006710.

PR 16-JUN-1989; US-367454.

PR 02-MAR-1992; US-847562.

PA (SCRI ) SCRIPPS RES INST.

PI Tobias PS, Ulevitch RJ;

DR WPI; 94-150499/18.

DR N-PSDB; Q63542.

PT Antibodies immunoreactive with lapine lipopolysaccharide binding

PT protein - useful for determn. of presence and amt. of human and

PT rabbit LBP.

PS Disclosure; Fig 19; 44pp; English.

CC R53459 shows a human lipopolysaccharide binding protein (LBP).

CC LBP is a glycoprotein that binds lipopolysaccharide that can be

CC present in the bloodstream of an individual susceptible to infection

CC by LPS-secreting gram negative bacteria.

SQ Sequence 477 AA;

Query Match 93.2%; Score 41; DB 1; Length 477;

Best Local Similarity 88.9%; Pred. No. 1.1;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSFFKV 9

|||||||:

DB 117 KVRKSFFKL 125

RESULT 8

ID R62501 standard; Protein; 481 AA.

AC R62501;

DT 03-MAY-1995 (first entry)

DE LBP amino terminal fragment, designated rLBP25.

KW Human; bactericidal/permeability-increasing protein; BPI; heparin;

KW binding agent; neutralisation; anti-coagulant effect; inhibition;

KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;

KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;

KW lipopolysaccharide; circulation; lymphokine; decontaminating; Helicobacter;

KW macrophage; activation; gastric ulcer; duodenal ulcer; antibiotic;

KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.

OS Homo sapiens.

PN Key

FT peptide 1. .25 Location/Qualifiers

FT /note= "signal peptide"

FT protein 26. .481

FT /note= "LBP N-terminal fragment"

FT WO9420532-A.

PD 15-SEP-1994.

PF 11-MAR-1994; U02465.

PR 12-MAR-1993; US-030644.

PR 15-JUL-1993; US-093202.

PR 14-JAN-1994; US-183222.

PA (XOMA ) XOMA CORP.

PI Little RG;

DR WPI; 94-302964/37.

DR N-PSDB; Q74302.

PT New human bactericidal permeability increasing peptides - derived

PT from the functional domains of BPI and having BPI activities such

PT as bactericidal activity

PS Disclosure; Page 151-154; 254pp; English.

CC This sequence represents the N-terminal fragment of lipopolysaccharide

CC binding protein (LBP) which has a molecular weight of approx 25 kD.

CC This protein is designated rLBP25 and may be used in conjunction with

CC the sequences given in R63682-750, R62087-100 and R62491-500 which are

CC peptides derived from human bactericidal/permeability-increasing

CC protein (BPI). BPI derived peptides such as these, may be used as

CC heparin binding agents, for neutralising the anti-coagulant effect of

CC heparin, for inhibiting angiogenesis, eg. associated with ocular

CC retinopathy, for inhibiting endothelial cell proliferation, for

CC contraception, for inhibiting malignant tumour cell proliferation,

CC for treating a chronic inflammatory disease state, eg. rheumatoid



CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 44; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFKV 9  
 |||||  
 Db 1 KVRKSFFKV 9

RESULT 3  
 R21842 R21842 standard; Protein; 454 AA.  
 AC R21842;  
 DT 22-JUL-1992 (first entry)  
 DE Sequence of chimera of lipopolysaccharide binding protein (LBP)  
 DE and BPI, a human 57kd protein which binds to the outer membrane of  
 DE susceptible gram negative bacteria  
 DE Endotoxin binding protein; diagnosis; therapy; endotoxemia;  
 KW Endotoxin shock.  
 OS Homo sapiens.

FH Key Location/Qualifiers  
 FT region 1. .197  
 FT /label= LBP  
 FT region 198. .454  
 FT /label= BPI

PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991; U05758.  
 PR 13-AUG-1990; US-567016.  
 PR 05-APR-1991; US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Maria MN, Scott RW;  
 DR WPI; 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Claim 4; Fig 23; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxemia and endotoxic shock.  
 SQ Sequence 454 AA;

Query Match 93.2%; Score 41; DB 1; Length 454;  
 Best Local Similarity 88.9%; Pred. No. 1.1;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFKV 9  
 |||||  
 Db 92 KVRKSFFKL 100

RESULT 4  
 R34140 R34140 standard; Protein; 481 AA.  
 AC R34140;  
 DT 20-JUL-1993 (first entry)  
 DE LBP-beta  
 KW Liposaccharide binding protein; immunostimulant; antibody;  
 KW immunoassay; LBP-alpha; recombinant.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT region 25. .481  
 FT /note= "mature protein"  
 FT peptide 1. .24  
 FT /note= "signal peptide"  
 PN WO9306228-A.  
 PD 01-APR-1993.  
 PF 28-SEP-1992; U08298.  
 PR 26-SEP-1991; US-765660.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Delegeane AM, Seilhamer JJ;  
 DR WPI; 93-117548/14.  
 DR N-PSDB; Q39305.  
 PT New form of lipo-saccharide binding protein, LBP-beta - useful as  
 PT immunostimulant, and antibodies are used to detect LBP-beta or  
 PT related proteins  
 PS Claim 1; Fig 2; 19pp; English.  
 CC PCR primers were used to amplify liposaccharide binding protein beta  
 CC encoding DNA from a commercially available cDNA prepn. derived from  
 CC human liver RNA by standard PCR. The amplified prods. were digested  
 CC with the corresp. restriction enzymes and isolated by PAGE. The  
 CC fragments were ligated into pMamNeo, a mammalian expression vector,  
 CC and cloned into E. coli. DNA from resulting clones was analysed and  
 CC sequenced. The sequence is similar to that of Schumann et al.,  
 CC Science (1990) 249: 1429-1431 but also contains an additional  
 CC sequence of 12 nucleotides in the coding region beginning after base  
 CC 826 of Genbank sequence 35533. The LBP-beta may be used to stimulate  
 CC the immune system and permits fine tuning of therapeutic protocols  
 CC designed to enhance immune function. Antibodies specifically reactive  
 CC with LBP-beta may be used in immunoassay systems for detection of  
 CC LBP-beta or related proteins. Antibodies which recognise epitopes  
 CC related to the novel portion of LBP-beta which distinguish it from  
 CC the previously disclosed human LBP are useful (there are 7 amino acid  
 CC substns. in LBP-beta relative to LBP-alpha).  
 SQ Sequence 481 AA;

Query Match 93.2%; Score 41; DB 1; Length 481;  
 Best Local Similarity 88.9%; Pred. No. 1.1;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVRKSFFKV 9  
 |||||  
 Db 117 KVRKSFFKL 125

RESULT 5  
 R41660 R41660 standard; Protein; 477 AA.  
 ID R41660;  
 AC R41660;  
 DT 18-MAR-1994 (first entry)  
 DE Human gram-negative lipopolysaccharide binding protein.  
 DE Gram negative bacteria; LPS; high density lipoprotein;  
 KW human; glycoprotein; gastrointestinal tract; endotoxaemia.  
 OS Homo sapiens.

FH Key Location/Qualifiers  
 FT protein 26. .477  
 FT /label= mature\_LBP  
 PN US5245013-A.  
 PD 14-SEP-1993.

PF 30-APR-1985; 728833.  
 PR 30-APR-1985; US-728833.  
 PR 30-DEC-1986; US-006710.  
 PR 16-JUN-1989; US-367454.  
 PA (ULEV/) ULEVITCH R.  
 PI Tobias P, Ulevitch R;  
 DR N-PSDB; Q47905.  
 DR WPI; 93-302704/38.

PT New glyco-protein which binds Gram negative lipo:polysaccharide -  
 PT isolated from acute phase sera, useful for determ. of endotoxin  
 PT in body fluid  
 PS Claim 1; Fig 19; 41pp; English.  
 CC A human liver cDNA library was prepared using Poly-A+ RNA isolated

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:49:54 ; Search time 135.78 Seconds  
(without alignments)  
1.570 Million cell updates/sec

Title: US-09-124-280A-23

Perfect score: 44  
Sequence: 1 KVRKSPFKV 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	44	100.0	9	1 R71794	Peptide neutralisi
2	44	100.0	9	1 W21611	Antibiotic potenti
3	41	93.2	454	1 R21842	Sequence of chimera
4	41	93.2	481	1 R34140	LBP-beta. New form
5	41	93.2	477	1 R41660	Human gram-negativ
6	41	93.2	197	1 R43887	Human lipopolysacc
7	41	93.2	477	1 R53459	Human lipopolysacc
8	41	93.2	481	1 R62501	LBP amino terminal
9	41	93.2	197	1 R67999	rLBP25 N-terminal
10	41	93.2	481	1 R68000	rLBP. Method for t
11	41	93.2	189	1 R68928	BPI(1-25)/lipopoly
12	41	93.2	198	1 R68930	LBP(1-146)/BPI(148
13	41	93.2	199	1 R68932	BPI(1-85)/lipopoly
14	41	93.2	196	1 R68934	BPI(1-85)/LBP(86-9
15	41	93.2	197	1 R68915	Lipopolysaccharide
16	41	93.2	197	1 R68914	Lipopolysaccharide
17	41	93.2	481	1 R68922	Lipopolysaccharide
18	41	93.2	18	1 R71986	Peptide inhibitor
19	41	93.2	15	1 R71987	Peptide inhibitor
20	41	93.2	15	1 R72045	Biologically activ
21	41	93.2	14	1 R72046	Biologically activ
22	41	93.2	14	1 R72049	Biologically activ
23	41	93.2	19	1 R72052	Biologically activ
24	41	93.2	18	1 R72053	Biologically activ
25	41	93.2	17	1 R72039	Biologically activ
26	41	93.2	17	1 R72043	Biologically activ
27	41	93.2	16	1 R72044	Biologically activ
28	41	93.2	18	1 R71989	Peptide inhibitor
29	41	93.2	15	1 R71991	Peptide inhibitor
30	41	93.2	18	1 R71992	Peptide inhibitor
31	41	93.2	481	1 R76601	Recombinant human
32	41	93.2	481	1 R76602	Recombinant human
33	41	93.2	481	1 R80014	Human LBP. Isolate
34	41	93.2	197	1 R80012	Human LBP(1-197).
35	41	93.2	485	1 R76242	Recombinant LBP. B
36	41	93.2	481	1 R81246	Recombinant lipopo
37	41	93.2	12	1 R90146	LBP-(90-101), pepti
38	41	93.2	14	1 R90150	Cys-bounded LBP-(9
39	41	93.2	481	1 R80994	Lipopolysaccharide
40	41	93.2	12	1 R95438	Endotoxin-neutrali
41	41	93.2	481	1 W05851	Recombinant lipopo
42	41	93.2	479	1 W16795	Recombinant endoto
43	41	93.2	481	1 W16830	Recombinant endoto

44 41 93.2 224 1 W16833 Recombinant endoto  
45 41 93.2 14 1 W20045 Human endotoxin-ne

ALIGNMENTS

RESULT 1  
R71794  
ID R71794 standard; peptide; 9 AA.  
AC R71794;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 25; Page 22; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which  
CC include units of formula: (A)n, where A is the cationic amino acid Lys  
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
CC Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides.  
SQ Sequence 9 AA;

Query Match 100.0%; Score 44; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVRKSPFKV 9  
DB 1 KVRKSPFKV 9

RESULT 2  
W21611  
ID W21611 standard; peptide; 9 AA.  
AC W21611;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #23.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN W09638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M; Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 27; Page 26; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to

**THIS PAGE BLANK (USPTO)**

Wed Sep 8 09:43:14 1999

SQ SEQUENCE 896 AA; 99504 MW; 8E7ED2CD CRC32;

Query Match 61.5%; Score 32; DB 11; Length 896;  
Best Local Similarity 54.5%; Pred. No. 1.8e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
| : : : : :  
Db 852 KDLVKKLPYQ 862

RESULT 15

Q63184 PRELIMINARY; PRT: 513 AA.  
AC Q63184;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE RNA-DEPENDENT INITIATION FACTOR-2 KINASE.  
GN EIF-2A.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;  
RX MEDLINE; 95035100.  
RA MELLOR H., FLOWERS K.M., KIMBALL S.R., JEFFERSON L.S.;  
RT "Cloning and characterization of a cDNA encoding rat PKR, the  
RT double-stranded RNA-dependent eukaryotic initiation factor-2  
RT kinase";  
RL Biochim. Biophys. Acta 1219:693-696(1994).  
DR EMBL; L29281; AAA61926.1; -  
DR PFAM; PF00035; dsrm; 2.  
DR PFAM; PF00069; pkinase; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW Initiation factor.  
SQ SEQUENCE 513 AA; 58258 MW; 5CBE316A CRC32;

Query Match 61.5%; Score 32; DB 11; Length 513;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSLSLKRLTY 10  
| : : : : :  
Db 260 KTYAIKRTY 269

Search completed: September 7, 1999, 22:47:29  
Job time: 7949 sec

DR EMBL; AF011333; AAC17636.1; -.  
DR PFAM; PF00040; fn2; 1.  
DR PFAM; PF00059; lectin\_c; 9.  
SQ SEQUENCE 1722 AA; 198270 MW; C2385581 CRC32;

Query Match 61.5%; Score 32; DB 4; Length 1722;  
Best Local Similarity 70.0%; Pred. No. 3.5e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSLKRLTY 10  
||| |  
Db 1516 KSKKLSRLTY 1525

RESULT 11

ID O75913 PRELIMINARY; PRT; 1722 AA.  
AC O75913;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE GP200-WR6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MCKAY P.E., RITTER M.A.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF064827; AAC62622.1; -.  
DR PFAM; PF00040; fn2; 1.  
DR PFAM; PF00059; lectin\_c; 9.  
SQ SEQUENCE 1722 AA; 198310 MW; 8F393829 CRC32;

Query Match 61.5%; Score 32; DB 4; Length 1722;  
Best Local Similarity 70.0%; Pred. No. 3.5e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSLKRLTY 10  
||| |  
Db 1516 KSKKLSRLTY 1525

RESULT 12

ID O49369 PRELIMINARY; PRT; 1086 AA.  
AC O49369;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE HYPOTHETICAL 122.7 KD PROTEIN.  
GN F106.150.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BEVAN M., WEICHELSELGARTNER M., FARTMANN B., GRANDERATH K., DAUNER D.,  
RA HERZL A., NEUMANN S., HOEISEL J., MEWES H.W., MAYER K.,  
RA SCHUELLER C.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL021811; CAAL16969.1; -.  
DR PFAM; PF01127; Sdh\_cyt; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 1086 AA; 122732 MW; 2A8B17D7 CRC32;

Query Match 61.5%; Score 32; DB 10; Length 1086;  
Best Local Similarity 60.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KSLSLKRLTY 10  
:|||||  
Db 668 QSLTKMKLTY 677

RESULT 13

ID O24203 PRELIMINARY; PRT; 87 AA.  
AC O24203;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
DE REVERSE TRANSCRIPTASE (FRAGMENT).  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC eudicotyledons; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
OC Poaceae; Oryza.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-NIPPONBARE; TRANSPOSON-TOS18;  
RX MEDLINE; 96353895.  
RA HIROCHIKA H., SUGIMOTO K., OTSUKI Y., TSUGAWA H., KANDA M.;  
RT "Retrotransposons of rice involved in mutations induced by tissue  
RT culture.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:7783-7788(1996).  
DR EMBL; D85877; BAA12903.1; -.  
KW RNA-directed DNA polymerase.  
FT NON\_TER 1  
FT NON\_TER 87  
SQ SEQUENCE 87 AA; 10345 MW; 9CCD2821 CRC32;

Query Match 61.5%; Score 32; DB 10; Length 87;  
Best Local Similarity 63.6%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
|||||  
Db 33 KSLRLKKSYYR 43

RESULT 14

ID O64146 PRELIMINARY; PRT; 896 AA.  
AC O64146;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE INTERLEUKIN-3 RECEPTOR BETA-SUBUNIT (FRAGMENT).  
GN RIL-3R<BETA>.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95370942.  
RA APPEL K., BUTTINI M., SAUTER A., GEBICKE-HARTER P.J.;  
RT "Cloning of rat interleukin-3 receptor beta-subunit from cultured  
RT microglia and its mRNA expression in vivo.";  
RL J. Neurosci. 15:5800-5809(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-WISTAR; TISSUE-BRAIN;  
RA GEBICKE-HARTER P.J.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; S79263; AAB35068.1; -.  
DR EMBL; AJ000555; CAA04186.1; -.  
DR PFAM; PF00041; fn3; 2.  
KW Signal.  
FT NON\_TER 1  
FT NON\_TER 896

RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
 RA JIWANI N., CARUSO A., BUSH D., SAFER H., PAYWELL D., PRABHAKAR S.,  
 RA MCDUGALL S., SHIMER G., GOVAL A., PIETROVSKI S., CHURCH G.M.,  
 RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.,  
 RT "Complete genome sequence of *Mechanobacterium thermoautotrophicum*  
 deltaH: functional analysis and comparative genomics.",  
 RL J. Bacteriol. 179:7135-7155(1997).  
 DR EMBL: AF000860; AAB85333.1; -.  
 SQ SEQUENCE 324 AA; 35553 MW; A836218C CRC32;

Query Match 63.5%; Score 33; DB 1; Length 324;  
 Best Local Similarity 63.6%; Pred. No. 42;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
 Db 112 RTCSDKRLTYR 122

RESULT 7  
 ID Q86795 PRELIMINARY; PRT; 524 AA.  
 AC Q86795;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE GLYCOPROTEIN G.  
 GN G.  
 OS rabies virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Rhabdoviridae; Lyssavirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94269833.  
 RA SAKAMOTO S., IDE T., NAKATAKE H., TOKIYOSHI S., YAMAMOTO M., KAWAI A.,  
 RA SMITH J.S.;  
 RT "Studies on the antigenicity and nucleotide sequence of the rabies  
 RT virus Nishigahara strain, a current seed strain used for dog vaccine  
 RT production in Japan."  
 RL Virus Genes 8:35-46(1994).  
 DR EMBL; S72465; AAB30931.1; -.  
 DR PFAM; PF00974; Rhabd\_glycop; 1.  
 SQ SEQUENCE 524 AA; 58472 MW; 4F6F6CC2 CRC32;

Query Match 63.5%; Score 33; DB 12; Length 524;  
 Best Local Similarity 60.0%; Pred. No. 67;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 10  
 Db 313 KSVSPRLSY 322

RESULT 8  
 ID Q49739 PRELIMINARY; PRT; 577 AA.  
 AC Q49739;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 61.3 KD PROTEIN B1620\_F2\_68.  
 GN B1620\_F2\_68.  
 OS *Mycobacterium leprae*.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA SMITH D.R., ROBISON K.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 DR EMBL; U00015; AAC43243.1; -.

DR PFAM; PF00990; DUF9; 1.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 222 242 POTENTIAL.  
 FT TRANSMEM 278 298 POTENTIAL.  
 FT TRANSMEM 314 334 POTENTIAL.  
 FT TRANSMEM 336 356 POTENTIAL.  
 FT TRANSMEM 358 378 POTENTIAL.  
 FT TRANSMEM 382 402 POTENTIAL.  
 FT TRANSMEM 420 440 POTENTIAL.  
 FT TRANSMEM 469 489 POTENTIAL.  
 FT TRANSMEM 502 522 POTENTIAL.  
 FT TRANSMEM 528 548 POTENTIAL.  
 FT TRANSMEM 549 569 POTENTIAL.  
 SQ SEQUENCE 577 AA; 61270 MW; 9C1A0C49 CRC32;

Query Match 61.5%; Score 32; DB 2; Length 577;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSLKRLTY 10  
 Db 107 LSIRLTY 114

RESULT 9  
 ID O13450 PRELIMINARY; PRT; 1953 AA.  
 AC O13450;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
 DE 'PPF3P'.  
 GN PPF3.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180;  
 RA YORIHUZI T., OHSUMI Y.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D38411; BAA22512.1; -.  
 SQ SEQUENCE 1953 AA; 219671 MW; EE9C015F CRC32;

Query Match 61.5%; Score 32; DB 3; Length 1953;  
 Best Local Similarity 87.5%; Pred. No. 3.9e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSLSLKRL 8  
 Db 964 KSLALKRL 971

RESULT 10  
 ID O60449 PRELIMINARY; PRT; 1722 AA.  
 AC O60449;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE DEC-205.  
 OS *Homo sapiens* (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98221117.  
 RA KATO M., NEIL T.K., CLARK G.J., MORRIS C.M., SORG R.V., HART D.N.J.;  
 RT "cDNA cloning of human DEC-205, a putative antigen-uptake receptor on  
 RT dendritic cells."  
 RL Immunogenetics 47:442-450(1998).

Query Match 65.4%; Score 34; DB 3; Length 809;  
 Best Local Similarity 70.0%; Pred. No. 63;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SLSLRLTYR 11  
 ||||:| |  
 Db 5 SLSLRQLDYR 14

RESULT 3  
 ID O23360 PRELIMINARY; PRT; 555 AA.  
 AC O23360;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 64.2 KD PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA EU ARABIDOPSIS SEQUENCING PROJECT, ESSA;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98121113.  
 RA BEVAN M., BANCROFT I., BENT E., LOVE K., GOODMAN H., DEAN C.,  
 RA BERGMAN R., DIRKSE W., VAN STAVEREN M., STIEREMA W., DROST L.,  
 RA RIDLEY P., HUDSON S.A., PATEL K., MURPHY G., PIFFANELLI P., WEDLER H.,  
 RA WEDLER E., WAMBUIT R., WEITZENEGGER T., POHL T.M., TERRY N.,  
 RA GIELEN J., VILLARROEL R., DE CLERCK R., VAN MONTAGU M., LECHARNY A.,  
 RA AUBORG S., GY I., KREIS M., LAO N., KAVANAGH T., HEMPEL S., KOTTER P.,  
 RA ANTIAN K.D., RIEGER M., SCHAEFER M., FUNK B., MUELLER-AUER S.,  
 RA SILVEY M., JAMES R., MONTFORT A., PONS A., PUIGDOMENECH P., DOUKA A.,  
 RA VOULELATOU E., MILIONI D., HATZIOPOULOS P., PIRAVANDI E., OBERMAIER B.,  
 RA HILBERT H., DUESTERHOFT A., MOORES T., JONES J.D.G., ENEVA T.,  
 RA PALME K., BENES V., RECHMAN S., ANSORGE M., COOKE R., BERGER C.,  
 RA DELSENY M., VOET M., VOLCKAERT G., MEWES H.W., KLOSTERMAN S.,  
 RA SCHUELLER C., CHALWATZIS N.;  
 RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of  
 Arabidopsis thaliana."  
 RL Nature 391:485-488(1998).  
 DR EMBL; 297337; CABL0285.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 555 AA; 64199 MW; 4DD3D266 CRC32;

Query Match 65.4%; Score 34; DB 10; Length 555;  
 Best Local Similarity 63.6%; Pred. No. 44;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSLSLRLTYR 11  
 | | | | |  
 Db 360 KYLLKRVYTK 370

RESULT 4  
 ID O10378 PRELIMINARY; PRT; 1967 AA.  
 AC O10378;  
 DT 01-JUL-1997 (TRENBLrel. 04, Created)  
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)  
 DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
 DE LARGE PROTEIN.  
 GN L.  
 OS rice transitory yellowing virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Rhabdoviridae; unclassified Rhabdoviridae.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA FANG R., LUO Z., ZHAO H.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB011257; BAA25160.1; -  
 SQ SEQUENCE 1967 AA; 223601 MW; 932E4E12 CRC32;

Query Match 65.4%; Score 34; DB 12; Length 1967;  
 Best Local Similarity 54.5%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSLSLRLTYR 11  
 | | | | |  
 Db 1056 KTATLKLRTK 1066

RESULT 5  
 ID Q38130 PRELIMINARY; PRT; 385 AA.  
 AC Q38130;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE ORF44.  
 OS Bacteriophage rlt.  
 OC Viruses.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96332668.  
 RA NAUTA A., VAN SINDEN D., KARSENS H., SMIT E., VENEMA G., KOK J.;  
 RT "Inducible gene expression mediated by a repressor-operator system  
 RT isolated from Lactococcus lactis bacteriophage rlt.";  
 Mol. Microbiol. 19:1331-1341(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96332669.  
 RA VAN SINDEN D., KARSENS H., KOK J., TERPSTRA P., RUITERS M.H.,  
 RA VENEMA G., NAUTA A.;  
 RT "Sequence analysis and molecular characterization of the temperate  
 RT lactococcal bacteriophage rlt.";  
 Mol. Microbiol. 19:1343-1355(1996).  
 DR EMBL; U38906; AAB18719.1; -  
 SQ SEQUENCE 385 AA; 43896 MW; 346F992B CRC32;

Query Match 64.4%; Score 33.5; DB 9; Length 385;  
 Best Local Similarity 64.3%; Pred. No. 39;  
 Matches 9; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 KSL---SLKRLTYR 11  
 | | | | |  
 Db 104 KSLRAKSLKRIYR 117

RESULT 6  
 ID O26923 PRELIMINARY; PRT; 324 AA.  
 AC O26923;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
 DE CONSERVED PROTEIN.  
 GN MTH835.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DELTA H;  
 RX MEDLINE; 98037514.  
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
 RA ADREGE T., BASHIRZADEH R., BLAKEY D., COOK R., GILBERT K.,  
 RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:26 ; search time 160.2 seconds  
(without alignments)  
4.226 Million cell updates/sec

Title: US-09-124-280A-22

Perfect score: 52

Sequence: 1 KSLSLKRLTYR 11

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL 10.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_nhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	69.2	128	11	O88955 rattus norv
2	34	65.4	809	3	O94527 schizosacch
3	34	65.4	555	10	O23360 arabidopsis
4	34	65.4	1967	12	O10378 rice transi
5	33.5	64.4	385	9	Q38130 bacterioph
6	33	63.5	324	1	O26923 methanobact
7	33	63.5	524	12	O86795 rabies viru
8	32	61.5	577	2	O49739 mycobacteri
9	32	61.5	1953	3	O13450 saccharomyc
10	32	61.5	1722	4	O60449 homo sapien
11	32	61.5	1722	4	O75913 homo sapien
12	32	61.5	1086	10	O49369 arabidopsis
13	32	61.5	87	10	O24203 oryza sativ
14	32	61.5	896	11	O64146 rattus norv
15	32	61.5	513	11	O63184 rattus norv
16	32	61.5	247	12	P89074 rabies viru
17	31.5	60.6	648	3	O74393 schizosacch
18	31.5	60.6	1180	12	O84919 pepper ring
19	31.5	60.6	1698	12	O84918 pepper ring
20	31	59.6	537	1	O28246 archaeglob
21	31	59.6	309	1	O58296 pyrococci
22	31	59.6	413	2	O25624 helicobacte
23	31	59.6	474	2	O44591 alcaligenes
24	31	59.6	99	2	Q52248 enterococu
25	31	59.6	717	2	P73701 synechocyst
26	31	59.6	431	2	Q92KN0 helicobacte
27	31	59.6	287	3	O14309 schizosacch
28	31	59.6	552	4	O92819 homo sapien
29	31	59.6	839	4	O00206 homo sapien

## ALIGNMENTS

## RESULT 1

O88955 PRELIMINARY; PRT; 128 AA.  
AC O88955;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE MONOCYTE DIFFERENTIATION ANTIGEN CD14 (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE DAWLEY; TISSUE-TESTIS;  
RA LIU S., SHAPIRO R.A.; BILLIAR T.R.;  
RT "Rat CD14 promoter."  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF087944; AAC35372.1; -  
FT NON\_TER 128 128  
SQ SEQUENCE 128 AA; 14265 MW; C6A183C8 CRC32;

Query Match 69.2%; Score 36; DB 11; Length 128;  
Best Local Similarity 72.7%; Pred. No. 4;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSLSLKRLTYR 11  
Db 89 RSLPLKRLTVR 99

## RESULT 2

O94527 PRELIMINARY; PRT; 809 AA.  
ID O94527  
AC O94527;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE GUANINE NUCLEOTIDE BINDING PROTEIN BETA SUBUNIT-LIKE.  
GN SPAC609.03.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H;  
RA SEEGER K., HARRIS D., WOOD V., RAJANDREAM M.A.; BARRELL B.G.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL035226; CAA22832.1; -  
SQ SEQUENCE 809 AA; 92498 MW; 61CA73DF CRC32;

30 31 59.6 225 5 Q17919  
31 31 59.6 441 5 O17190  
32 31 59.6 405 5 O15921  
33 31 59.6 428 5 Q24386  
34 31 59.6 657 5 O62361  
35 31 59.6 244 5 O86231  
36 31 59.6 127 6 Q04209  
37 31 59.6 552 6 O97711  
38 31 59.6 428 10 O24457  
39 31 59.6 321 10 Q41413  
40 31 59.6 305 10 Q41416  
41 31 59.6 585 10 O81444  
42 31 59.6 439 10 Q92TK5  
43 31 59.6 552 11 O35776  
44 31 59.6 473 11 Q05701  
45 31 59.6 552 11 P70312

Q17919 caenorhabdi  
O17190 caenorhabdi  
O15921 trypanosoma  
Q24386 drosophila  
O62361 caenorhabdi  
O86231 plasmodium  
Q04209 ovis aries  
O97711 bos taurus  
O24457 arabidopsis  
Q41413 solanum tub  
Q41416 solanum tub  
O81444 oryza sativ  
Q92TK5 catharanthu  
O35776 rattus norv  
Q05701 rattus norv  
P70312 mus musculu



**THIS PAGE BLANK (USPTO)**

Qy 1 KSLSLRLTY 10  
||: || ||:  
Db 809 KSIPKLNLF 818

Search completed: September 7, 1999, 23:59:01  
Job time: 513 sec

;  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-049-871-6

Query Match 100.0%; Score 58; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTKCKFLKKC 10  
Db 2 KTKCKFLKKC 11

RESULT 3  
US-07-819-893-2  
; Sequence 2, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514

;  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
; US-07-819-893-2

Query Match 100.0%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.003;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTKCKFLKKC 10  
Db 1 KTKCKFLKKC 10

RESULT 4  
US-07-819-893-6  
; Sequence 6, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
; US-07-819-893-6

Query Match 100.0%; Score 58; DB 1; Length 11;

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:48 ; Search time 82.37 Seconds  
(without alignments)  
1.198 Million cell updates/sec

Title: US-09-124-280A-31

Perfect score: 58  
Sequence: 1 KTKCKFLKCC 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9869381 residues

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	10	1	US-08-049-871-2
2	58	100.0	11	1	US-08-049-871-6
3	58	100.0	10	1	US-07-819-893-2
4	58	100.0	11	1	US-07-819-893-6
5	58	100.0	10	1	US-08-280-397-2
6	58	100.0	11	1	US-08-280-397-6
7	58	100.0	11	2	US-08-218-026-49
8	58	100.0	11	2	US-08-218-026-50
9	58	100.0	10	2	US-08-653-632-49
10	58	100.0	11	2	US-08-653-632-50
11	58	100.0	10	2	US-08-456-1128-31
12	58	100.0	11	2	US-08-456-1128-35
13	43	74.1	7	1	US-08-049-871-1
14	43	74.1	7	1	US-07-819-893-1
15	43	74.1	7	1	US-08-280-397-1
16	43	74.1	7	2	US-08-456-1128-30
17	37	63.8	10	1	US-08-097-830E-15
18	37	63.8	10	1	US-08-097-830E-17
19	37	63.8	10	2	US-08-456-112B-15
20	37	63.8	10	2	US-08-456-112B-17
21	35	60.3	2713	3	PCT-US96-01735-1
22	34	58.6	553	1	US-08-328-322-10
23	34	58.6	509	1	US-08-328-322-19
24	33	56.9	1003	1	US-07-991-867B-6
25	33	56.9	10	1	US-08-097-830E-16
26	33	56.9	1003	1	US-08-107-755A-6
27	33	56.9	223	1	US-08-430-633-1
28	33	56.9	10	2	US-08-456-112B-16
29	33	56.9	1824	2	US-08-680-327-3
30	33	56.9	391	2	US-08-644-034A-1
31	32	55.2	551	1	US-08-120-960-2
32	32	55.2	1089	1	US-08-180-195-36
33	32	55.2	1089	1	US-08-168-917-4
34	32	55.2	45	1	US-08-632-691-1
35	32	55.2	46	1	US-08-462-661A-13
36	32	55.2	45	1	US-08-462-661A-15
37	32	55.2	1089	2	US-08-477-329-36
38	32	55.2	1089	2	US-08-475-458-36
39	32	55.2	1089	2	US-08-460-510-4

40 32 55.2 1089 3 PCT-US92-00730-4 Sequence 4, Appli  
41 32 55.2 1089 3 PCT-US92-00862-4 Sequence 4, Appli  
42 31.5 54.3 10 1 US-08-097-830E-18 Sequence 18, Appli  
43 31.5 54.3 10 2 US-08-456-112B-18 Sequence 18, Appli  
44 31 53.4 98 1 US-07-623-611-6 Sequence 6, Appli  
45 31 53.4 551 3 PCT-US95-02058-9 Sequence 9, Appli

ALIGNMENTS

RESULT 1  
US-08-049-871-2  
; Sequence 2, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-2  
  
Query Match 100.0%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.003; 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0;  
  
QY 1 KTKCKFLKCC 10  
| | | | | | | | | |  
Db 1 KTKCKFLKCC 10  
  
RESULT 2  
US-08-049-871-6  
; Sequence 6, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the

CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 63.8%; Score 37; DB 1; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 2;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
 :|:|:|:|:  
 Db 1 RTRCRFKRRC 10

RESULT 15  
 W21603  
 ID W21603 standard; peptide: 10 AA.  
 AC W21603;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #15.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 OS permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide bond 4..10  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 19; Page 25; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 63.8%; Score 37; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 2;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
 | | | | |  
 Db 1 KFKCKFKKFC 10

Search completed: September 7, 1999, 22:50:03  
 Job time: 7884 sec

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 CC Sequence 7 AA;

Query Match 74.1%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CKFLKCC 10  
 | | | | |  
 Db 1 CKFLKCC 7

## RESULT 12

W38732  
 ID W38732 standard; Protein; 37 AA.  
 AC W38732;  
 DT 10-NOV-1998 (first entry)  
 DE Streptococcus pneumoniae protein of unknown function.  
 KW Streptococcus pneumoniae protein; genetic immunisation; antagonist;  
 KW immunological response; inoculation; antibody production; inhibitor;  
 KW T cell immune response; antimicrobial compound; bacterial adhesion;  
 KW extracellular matrix protein; protein-mediated cell invasion; wound;  
 KW pathogenesis.  
 OS Streptococcus pneumoniae.  
 PN W09743303-A1.  
 PD 20-NOV-1997.  
 PF 14-MAY-1997; U07950.  
 PR 14-MAY-1996; US-017670.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO,  
 PI Stodola RK;  
 DR WPI: 98-008793/01.  
 DR N-PSDB; 198774.  
 DR Novel Streptococcus pneumoniae proteins and related DNA - useful for  
 PT diagnosing anti-microbial agents for treatment of bacterial  
 PT infections  
 PS Claim 12: Pages 460-461; 483pp; English.  
 CC This sequence represents a Streptococcus pneumoniae protein of  
 CC unknown function, and is encoded by a DNA sequence of the invention.  
 CC The DNA sequences were isolated from Streptococcus pneumoniae strain  
 CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the  
 CC invention can be used to identify compounds which interact with and  
 CC inhibit or activate the activity of the proteins. Antagonists can be  
 CC used to treat diseases caused by S. pneumoniae proteins, through genetic  
 CC immunisation. They can also be used to induce an immunological response  
 CC in a mammal by inoculation with the S. pneumoniae proteins or delivery  
 CC of the encoding nucleic acids in a vector adequate to produce antibody  
 CC and/or T cell immune responses to protect the animal from disease. The  
 CC proteins can also be used to identify antimicrobial compounds which are  
 CC capable of inhibiting their bioactivity. In particular the proteins of  
 CC the invention can be used to prevent adhesion of bacteria to mammalian  
 CC extracellular matrix proteins on in-dwelling devices or in wounds, to  
 CC block protein-mediated mammalian cell invasion, and to block the normal  
 CC progression of pathogenesis in infections initiated other than by the  
 CC implantation of in-dwelling devices or other surgical techniques.  
 CC Sequence 37 AA;

Query Match 65.5%; Score 38; DB 1; Length 37;  
 Best Local Similarity 70.0%; Pred. No. 4.1;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10  
 : | | | | |  
 Db 20 RNKDKFLKCC 29

## RESULT 13

R71786  
 ID R71786 standard; peptide; 10 AA.  
 AC R71786;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide bond 4..10  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI: 95-075190/10.

PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 17: Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 CC Sequence 10 AA;

Query Match 63.8%; Score 37; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 2;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10  
 | | | | |  
 Db 1 KFKCKFKKCC 10

## RESULT 14

R71788  
 ID R71788 standard; peptide; 10 AA.  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide bond 4..10  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI: 95-075190/10.

PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 19: Page 21; 26pp; English.

CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic

CC acid sequence of a peptide with endotoxin neutralising activity combined  
 CC with at least 1 different amino acid sequence of a peptide that has  
 CC bactericidal activity; and (e) a prosthetic device which has a  
 CC sufficient amount of BP attached to the surface to inhibit bacterial  
 CC infection. The peptides of the invention are used in treating bacterial  
 CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7 to  
 CC 10-9 M, and Escherichia coli. The peptides are also used to treat  
 CC endotoxin shock. The present sequence represents a peptide derived  
 CC from human neutrophil granule bactericidal protein from an example of  
 CC the present invention.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 58; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00086;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTKCKFLKCC 10  
 |||||  
 Db 2 KTKCKFLKCC 11

RESULT 9  
 R33525  
 ID R33525 standard; peptide; 7 AA.  
 AC R33525;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT disulfide\_bond 1. .7  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 3; Page 32; 39pp; English.

CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His - Phe/Tyr/Trip - Leu/Ile/Val)n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 7 AA;

Query Match 74.1%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CKFLKCC 10  
 |||||  
 Db 1 CKFLKCC 7

RESULT 10  
 R39287  
 ID R39287 standard; peptide; 7 AA.  
 AC R39287;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT disulfide\_bond 1. .7  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-243143/30.

PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 3; Page 32; 45pp; English.

CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 7 AA;

Query Match 74.1%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CKFLKCC 10  
 |||||  
 Db 1 CKFLKCC 7

RESULT 11  
 W21618  
 ID W21618 standard; peptide; 7 AA.  
 AC W21618;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #30.  
 KW Potentiator; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT disulfide\_bond 1. .7  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 34; Page 27; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to

```

PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M;
DR WPI: 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 35; Page 27; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to
CC potentiate the activity of an antibiotic when they are co-administered
CC with the antibiotic. Compositions containing these peptides are used
CC to treat or prevent microbial infections. These peptides bind to
CC lipopolysaccharide on the bacteria so may increase permeability of
CC the outer bacterial membrane to the antibiotic, allowing a reduction
CC in the dose of antibiotic required by 10-90% of the normal dose for
CC in vivo or in vitro application. Any toxic side effects are
CC correspondingly reduced.
SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
DB 1 KTKCKFLKCC 10

RESULT 6
W21623
ID W21623 standard; peptide; 11 AA.
AC W21623;
DE 26-AUG-1997 (first entry)
KW Antibiotic potentiating peptide #35.
KW Potentiator; antibiotic; microbial infection; lipopolysaccharide;
KW permeability; outer bacterial membrane.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide_bond 5..11
PN WO9638163-A1.
PD 05-DEC-1996.
PF 29-MAY-1996; E02313.
PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M;
PI WPI: 97-034095/03.
DR Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 39; Page 28; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to
CC potentiate the activity of an antibiotic when they are co-administered
CC with the antibiotic. Compositions containing these peptides are used
CC to treat or prevent microbial infections. These peptides bind to
CC lipopolysaccharide on the bacteria so may increase permeability of
CC the outer bacterial membrane to the antibiotic, allowing a reduction
CC in the dose of antibiotic required by 10-90% of the normal dose for
CC in vivo or in vitro application. Any toxic side effects are
CC correspondingly reduced.
SQ Sequence 11 AA;

Query Match 100.0%; Score 58; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
DB 2 KTKCKFLKCC 11

RESULT 7
W62434
ID W62434 standard; peptide; 10 AA.

```

---

```

AC W62434;
DT 01-OCT-1998 (first entry)
DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:49.
KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;
KW bactericidal permeability increasing factor; B/PI; neutralising;
KW human neutrophil granule bactericidal protein.
OS Homo sapiens.
OS Synthetic.
PN US5786324-A.
PD 28-JUL-1998.
PF 24-MAR-1994; 218026.
PR 24-MAR-1994; US-218026.
PA (MINU) UNIV MINNESOTA.
PI Gray B, Haseman JR, Mayo K;
PI WPI: 98-436578/37.
DR Bactericidal and endotoxin-neutralising peptides - used in treating
PT e.g. Pseudomonas species infection and in protectively coating
PT prosthetic devices
PS Example 1; Column 18; 46pp; English.
CC The present invention describes bactericidal peptides (BP): (a) for
CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)
CC with both endotoxin neutralising activity and bactericidal activity for
CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising
CC activity and bactericidal activity, the peptide having at least 1 amino
CC acid sequence of a peptide with endotoxin neutralising activity combined
CC with at least 1 different amino acid sequence of a peptide that has
CC bactericidal activity; and (e) a prosthetic device which has a
CC sufficient amount of BP attached to the surface to inhibit bacterial
CC growth. The peptides of the invention are used in treating bacterial
CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7 to
CC 10-9 M, and Escherichia coli. The peptides are also used to treat
CC endotoxin shock. The present sequence represents a peptide derived of
CC from human neutrophil granule bactericidal protein from an example of
CC the present invention.
SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
DB 1 KTKCKFLKCC 10

RESULT 8
W62435
ID W62435 standard; peptide; 11 AA.
AC W62435;
DT 01-OCT-1998 (first entry)
DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:50.
KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;
KW bactericidal permeability increasing factor; B/PI; neutralising;
KW human neutrophil granule bactericidal protein.
OS Homo sapiens.
OS Synthetic.
PN US5786324-A.
PD 28-JUL-1998.
PF 24-MAR-1994; 218026.
PR 24-MAR-1994; US-218026.
PA (MINU) UNIV MINNESOTA.
PI Gray B, Haseman JR, Mayo K;
PI WPI: 98-436578/37.
DR Bactericidal and endotoxin-neutralising peptides - used in treating
PT e.g. Pseudomonas species infection and in protectively coating
PT prosthetic devices
PS Example 1; Column 18; 46pp; English.
CC The present invention describes bactericidal peptides (BP): (a) for
CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)
CC with both endotoxin neutralising activity and bactericidal activity for
CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising
CC activity and bactericidal activity, the peptide having at least 1 amino

```



PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 DR WPI: 93-094304/11.  
 DT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PS comprising sequences retro-orientated  
 PS Claim 8; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula 8-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood;  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 58; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00086;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
 | | | | | | | |  
 DB 2 KTKCKFLKKC 11

RESULT 3  
 ID R39288 standard; peptide; 10 AA.  
 AC R39288:  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-243143/30.  
 DT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 4; Page 32; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins; and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0008;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
 | | | | | | | |  
 DB 1 KTKCKFLKKC 10

RESULT 4  
 ID R39292 standard; peptide; 11 AA.  
 AC R39292:  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 5..11  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-243143/30.  
 DT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 8; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins; and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 58; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00086;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
 | | | | | | | |  
 DB 2 KTKCKFLKKC 11

RESULT 5  
 ID W21619 standard; peptide; 10 AA.  
 AC W21619:  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #31.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:02; Search time 135.78 Seconds  
(without alignments)  
1.744 Million cell updates/sec

Title: US-09-124-280A-31

Perfect score: 58

Sequence: 1 KTKCKFLKCC 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	10	1 R33526	Peptide for treati
2	58	100.0	10	1 R33531	Peptide for treati
3	58	100.0	10	1 R39288	Endotoxin lipid A
4	58	100.0	10	1 R39292	Endotoxin lipid A
5	58	100.0	10	1 W21619	Antibiotic potenti
6	58	100.0	10	1 W21623	Antibiotic potenti
7	58	100.0	10	1 W62434	Human neutrophil g
8	58	100.0	10	1 W62435	Human neutrophil g
9	43	74.1	7	1 R33525	Peptide for treati
10	43	74.1	7	1 R39287	Endotoxin lipid A
11	43	74.1	7	1 W21618	Antibiotic potenti
12	38	65.5	37	1 W38732	Streptococcus pneu
13	37	63.8	10	1 R71786	Peptide neutralisi
14	37	63.8	10	1 R71788	Peptide neutralisi
15	37	63.8	10	1 W21603	Antibiotic potenti
16	37	63.8	10	1 W21605	Antibiotic potenti
17	35	60.3	2749	1 R13887	Inositol-3-phospha
18	35	60.3	2713	1 W00168	Human T cell inosi
19	35	60.3	2695	1 W07632	Human type I inosi
20	35	60.3	171	1 W57201	Human interleukin-
21	35	60.3	18	1 W99423	Cationic peptide B
22	35	60.3	18	1 W99424	Cationic peptide B
23	34	58.6	509	1 R96086	Yeast calcineurin
24	34	58.6	263	1 W01573	Protein encoded by
25	34	58.6	263	1 W40374	Human breast cance
26	34	58.6	579	1 W40378	Human breast cance
27	34	58.6	173	1 Y07084	Renal cancer assoc
28	34	58.6	312	1 Y07052	Renal cancer assoc
29	33	56.9	1003	1 R29648	AmEPV Spheroidin P
30	33	56.9	1003	1 R55576	AmEPV spheroidin.
31	33	56.9	383	1 R62083	Lactococcus lactis
32	33	56.9	10	1 R71787	Peptide neutralisi
33	33	56.9	223	1 R87020	B2LF2 protein. Eps
34	33	56.9	88	1 R30991	Mouse Mat-8 polype
35	33	56.9	10	1 W21604	Antibiotic potenti
36	33	56.9	1003	1 W41301	AmEPV entomopoxvir
37	33	56.9	223	1 W47351	Epstein-barr virus
38	33	56.9	1822	1 W48360	Tomato Prf protein
39	33	56.9	137	1 W38577	S. pneumoniae 50S
40	33	56.9	585	1 W97757	S-region transfer
41	32.5	56.0	99	1 R04135	Tat gene of simian
42	32	55.2	1089	1 R06910	Alpha type PDGF re
43	32	55.2	1089	1 R08267	Platelet derived g

#### ALIGNMENTS

```

RESULT 1
R33526
ID R33526 standard; peptide; 10 AA.
AC R33526;
DT 07-JUL-1993 (first entry)
DE Peptide for treating septic shock.
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;
KW cytokine release control; treatment; pertussis; bacterial meningitis;
KW HIV related infections; polymyxin B.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide_bond 4..10
PN ZA9200943-A.
PD 25-NOV-1992.
PE 10-FEB-1992: 000943.
PR 11-FEB-1991: US-658744.
PA (PORR/) PORR M.
PI Porro M;
DR WPI; 93-094304/11.
PT New peptide for treatment or prevention of toxic shock - comprises
PT specified sequences of aminoacid(s) and analogs
PT comprising sequences retro-orientated
PS Claim 4; Page 32; 39pp; English.
CC This is a specific example of a generic peptide of formula
CC R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where n = 1-100
CC and each R is H, an amino acid residue or a fatty acid residue.
CC The peptide is useful for treating or preventing septic shock,
CC mixing with polymyxin B to reduce its toxicity; removing
CC endotoxins from blood, sera or other fluids (in vivo or in
CC vitro); controlling release of cytokines induced by endotoxins;
CC as diagnostic reagents to detect and quantify toxins in blood
CC or sera; preparing non-toxic antigenic complexes of lipid A or
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial
CC meningitis and HIV-related infections. The usual dose is 10-100
CC ug/kg/day, given parenterally. It binds to the same sites as
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It
CC has no antibiotic activity; does not lyse erythrocytes; has no
CC toxicity in mice when injected at 50mg/kg and is relatively unstable
CC against proteases.
SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
   |||||
Db 1 KTKCKFLKCC 10

RESULT 2
R33531
ID R33531 standard; peptide; 11 AA.
AC R33531;
DT 07-JUL-1993 (first entry)
DE Peptide for treating septic shock.
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;
KW cytokine release control; treatment; pertussis; bacterial meningitis;
KW HIV related infections; polymyxin B.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide_bond 5..11
PN ZA9200943-A.
PD 25-NOV-1992.
PE 10-FEB-1992: 000943.

```

Platelet aggregati  
Type B human plate



OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TASSONE F., VILLARD L., GARDINER K.;  
 RT "Sequence, genomic organization and map localization of the human SR  
 RL protein gene rA4.";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF023142; AAD09327.1; -  
 FT NON\_TER 1  
 FT 1  
 SQ SEQUENCE 1157 AA; 127040 MW; 107FE6DF CRC32;

Query Match 74.4%; Score 32; DB 4; Length 1157;  
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KFLKKC 7  
 ||: |||  
 Db 59 KFIKKC 64

RESULT 15  
 OI7928 PRELIMINARY; PRT; 338 AA.  
 AC OI7928;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE NHR-52 PROTEIN.  
 GN NHR-52.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LLOYD C.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.; III of C.  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; 263233; CAB05761.1; -  
 SQ SEQUENCE 338 AA; 39582 MW; 0F9FB43C CRC32;

Query Match 74.4%; Score 32; DB 5; Length 338;  
 Best Local Similarity 71.4%; Pred. No. 57;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CKFLKKC 7  
 ||: |||  
 Db 39 CKYDKKC 45

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DE 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 GN HD-21P PROTEIN.  
 OS THOM1.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Solanum.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SHOOT APEX;  
 RX MEDLINE; 95218609.  
 RA METSSNER R., THERES K.;  
 RT "Isolation and characterization of the tomato homeobox gene THOM1.";  
 RL Planta 195;541-547(1995).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; X91212; CAA62608.1; -.  
 DR MENDEL; 8924; Lyces:1345;1.  
 DR PFAM; PF00046; homeobox; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 SQ SEQUENCE 286 AA; 32143 MW; 7C2E2DEC CRC32;

Query Match 76.7%; Score 33; DB 10; Length 286;  
 Best Local Similarity 57.1%; Pred. No. 33;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFLKRC 7  
 |::||:|  
 DB 188 CEYLKRC 194

RESULT 11  
 Q42975 ID O42975 PRELIMINARY; PRT; 972 AA.  
 AC O42975;  
 DT 01-JAN-1999 (TRENBLrel. 09, Created)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last sequence update)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
 DE HYPOTHETICAL 113.2 KD PROTEIN C20F10.05 IN CHROMOSOME II.  
 GN SPBC20F10.05.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA LYNE M., MCDUGALL R., RAJANDREAM M.A., BARRELL B.G., BECK A.,  
 RA REINHARDT R.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL021747; CAA16845.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 972 AA; 113248 MW; 646695C1 CRC32;

Query Match 74.4%; Score 32; DB 3; Length 972;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CRFLKRC 7  
 |::||:|  
 DB 338 CDFLKNC 344

RESULT 12  
 Q14660 ID Q14660 PRELIMINARY; PRT; 2713 AA.  
 AC Q14660;

DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE INOSITOL 1,4,5-TRISPHOSPHATE RECEPTOR TYPE 1.  
 GN ITPR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-T LYMPHOCYTE;  
 RX MEDLINE; 95155356.  
 RA HARNICK D.J., JAYARAMAN T., MA Y., MULIERI P., GO L.O., MARKS A.R.;  
 RT "The human type 1 inositol 1,4,5-trisphosphate receptor from T  
 lymphocytes. Structure, localization, and tyrosine phosphorylation.";  
 RL J. Biol. Chem. 270:2833-2840(1995).  
 DR EMBL; L38019; AAB04947.1; -.  
 DR PFAM; PF01365; RYDR\_ITPR; 1.  
 SQ SEQUENCE 2713 AA; 308697 MW; A16B46FB CRC32;

Query Match 74.4%; Score 32; DB 4; Length 2713;  
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKRC 7  
 |::||:|

DB 1323 KFIKRC 1328

RESULT 13  
 Q14643 ID Q14643 PRELIMINARY; PRT; 2695 AA.  
 AC Q14643;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE HUMAN TYPE 1 INOSITOL 1,4,5-TRISPHOSPHATE RECEPTOR.  
 GN INSP3R1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-UTERUS, AND MYELOID;  
 RX MEDLINE; 95031918.  
 RA YAMADA N., MAKINO Y., CLARK R.A., PEARSON D.W., MATTEI M.G.,  
 RA GUNNET J.L., OHAMA E., FUJINO I., MIYAWAKI A., FURUICHI T.,  
 RA MIKOSHIBA K.;  
 RT "Human inositol 1,4,5-trisphosphate type-1 receptor, InSP3R1:  
 structure, function, regulation of expression and chromosomal  
 localization.";  
 RL Biochem. J. 302:781-790(1994).  
 DR EMBL; D26070; BAA05065.1; -.  
 DR PFAM; PF01365; RYDR\_ITPR; 1.  
 SQ SEQUENCE 2695 AA; 306769 MW; 99197877 CRC32;

Query Match 74.4%; Score 32; DB 4; Length 2695;  
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKRC 7  
 |::||:|

DB 1307 KFIKRC 1312

RESULT 14  
 Q95104 ID Q95104 PRELIMINARY; PRT; 1157 AA.  
 AC Q95104;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE PRE-MRNA SPLICING SR PROTEIN RA4 (FRAGMENT).

OC core eudicots; Asteridae; euasterids II; Apiales; Apiaceae;  
 OC Pimpinella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE; 98196666  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA MOON Y.H., CHOI S.B., KIM J.I., KIM J.C., HAN T.J., CHO S.H.,  
 RA LEE K.W.;  
 RL Mol. Cells 6:697-703(1996).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; X94449; CAA64221.1; -.  
 DR MENDEL; 12128; PIMB; 1345; 3.  
 DR PFAM; PF00046; homeobox; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR Homeobox; DNA-binding; Nuclear protein; Homeobox.  
 KW Homeobox 302 AA; 33664 MW; A8699724 CRC32;  
 SQ SEQUENCE 302 AA; 33664 MW; A8699724 CRC32;

Query Match 83.7%; Score 36; DB 10; Length 302;  
 Best Local Similarity 71.4%; Pred. No. 9.2;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKCC 7

I:||||:|

Db 200 CEFLKRC 206

RESULT 7

Q43441 ID Q43441 PRELIMINARY; PRT; 171 AA.  
 AC Q43441;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-WAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE HOMEBOX-LEUCINE ZIPPER PROTEIN (FRAGMENT).  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 OC Glycine.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. WHANG-KEUM; TISSUE-GLOBULAR SOMATIC EMBRYO;  
 RA MOON Y.H., CHOI S.B., LEE K.W.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; U30475; AAA74017.1; -.  
 DR MENDEL; 8292; GLIma; 1345; 2.  
 DR PFAM; PF00046; homeobox; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 171 AA; 19903 MW; 261D385C CRC32;

Query Match 83.7%; Score 36; DB 10; Length 171;  
 Best Local Similarity 71.4%; Pred. No. 6;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKCC 7

I:||||:|

Db 62 CEFLKRC 68

RESULT 8

O67915 ID O67915 PRELIMINARY; PRT; 189 AA.  
 AC O67915;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 22.1 KD PROTEIN.  
 GN AQ\_2184.  
 OS Aquifex aeolicus.

OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE; 98196666  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus";  
 RL Nature 392:353-358(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE000776; AAC07887.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 189 AA; 22088 MW; 73CF88CE CRC32;

Query Match 79.1%; Score 34; DB 2; Length 189;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKCC 7

|||||

Db 131 KFLKCC 136

RESULT 9

Q39927 ID Q39927 PRELIMINARY; PRT; 236 AA.  
 AC Q39927;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE HOMEODOMAIN PROTEIN.  
 OS Helianthus annuus (Common sunflower).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids II; Asterales; Asteraceae;  
 OC Helianthus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-ROOT;  
 RA GONZALEZ D.H., VALLE E.M., GAGO G., CHAN R.L.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; L48485; AAA79778.1; -.  
 DR MENDEL; 8447; HELan; 1345; 1.  
 DR PFAM; PF00046; homeobox; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 SQ SEQUENCE 236 AA; 26142 MW; 188C4BB0 CRC32;

Query Match 76.7%; Score 33; DB 10; Length 236;  
 Best Local Similarity 57.1%; Pred. No. 28;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKCC 7

I:||||:|

Db 156 CEYLKRC 162

RESULT 10

Q42437 ID Q42437 PRELIMINARY; PRT; 286 AA.  
 AC Q42437;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
 DE 01-JUN-1998 (TRENBLrel. 06, Last annotation update)  
 DE HYPOTHETICAL 28.0 KD PROTEIN.  
 OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).  
 OC Viruses: dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.  
 RN [1]  
 RA SEQUENCE FROM N.A.  
 RP BAHR U., TIDONA C.A., DARAI G.;  
 RL VIRUS GENES 0:0-0(1997).  
 DR EMBL: AF003534; AAB94452.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 239 AA; 27979 MW; 7B15C915 CRC32;

Query Match 86.0%; Score 37; DB 12; Length 239;  
 Best Local Similarity 71.4%; Pred. No. 5;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
 ||:||||  
 Db 87 CKYVKCC 93

RESULT 3  
 Q39862  
 ID Q39862 PRELIMINARY; PRT; 284 AA.  
 AC Q39862;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE HOMEBOX-LEUCINE ZIPPER PROTEIN.  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 OC Glycine.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. HWANG-KEUM;  
 RA MOON Y.H., CHOI S.B., KIM J.I., HAN T.J., CHO S.H., KIM W.T.,  
 RA LEE K.W.;  
 RL MOL. Cells 6:366-373(1996).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: X92489; CAA63222.1; -.  
 DR MENDEL: 8291; GLYMa.1345.1;  
 DR PFAM: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Homeobox.  
 SQ SEQUENCE 284 AA; 31919 MW; 9D3F8C00 CRC32;

Query Match 83.7%; Score 36; DB 10; Length 284;  
 Best Local Similarity 71.4%; Pred. No. 8.8;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
 ||:||||  
 Db 175 CEFLKRC 181

RESULT 4  
 Q40780  
 ID Q40780 PRELIMINARY; PRT; 319 AA.  
 AC Q40780;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE HOMEBOX-LEUCINE ZIPPER PROTEIN.  
 OS Pimpinella brachycarpa.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids II; Apiales; Apiaceae;  
 OC Pimpinella.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SHOOT APEX;  
 RA MOON Y.H., CHOI S.B., KIM J.I., KIM J.C., HAN T.J., CHO S.H.,  
 RA LEE K.W.;  
 RL MOL. Cells 6:697-703(1996).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: X94375; CAA64152.1; -.  
 DR MENDEL: 12126; PIMbr.1345.1.  
 DR PFAM: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Homeobox.  
 SQ SEQUENCE 319 AA; 36055 MW; D71EE66A CRC32;

Query Match 83.7%; Score 36; DB 10; Length 319;  
 Best Local Similarity 71.4%; Pred. No. 9.6;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
 ||:||||  
 Db 214 CEFLKRC 220

RESULT 5  
 Q40781  
 ID Q40781 PRELIMINARY; PRT; 318 AA.  
 AC Q40781;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE HOMEBOX-LEUCINE ZIPPER PROTEIN.  
 OS Pimpinella brachycarpa.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids II; Apiales; Apiaceae;  
 OC Pimpinella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SHOOT APEX;  
 RA MOON Y.H., CHOI S.B., KIM J.I., KIM J.C., HAN T.J., CHO S.H.,  
 RA LEE K.W.;  
 RL MOL. Cells 6:697-703(1996).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: X95193; CAA64491.1; -.  
 DR MENDEL: 12127; PIMbr.1345.2.  
 DR PFAM: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Homeobox.  
 SQ SEQUENCE 318 AA; 35936 MW; CB5B0D6F CRC32;

Query Match 83.7%; Score 36; DB 10; Length 318;  
 Best Local Similarity 71.4%; Pred. No. 9.6;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
 ||:||||  
 Db 213 CEFLKRC 219

RESULT 6  
 Q40782  
 ID Q40782 PRELIMINARY; PRT; 302 AA.  
 AC Q40782;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE HOMEBOX-LEUCINE ZIPPER PROTEIN.  
 OS Pimpinella brachycarpa.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC Pimpinella.

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:39 ; Search time 160.2 Seconds  
(without alignments)  
2.689 Million cell updates/sec

Title: US-09-124-280A-30

Perfect score: 43

Sequence: 1 CKFLKKC 7

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL10.\*  
1: sp.archaea.\*  
2: sp.bacteria.\*  
3: sp.fungi.\*  
4: sp.human.\*  
5: sp.invertebrate.\*  
6: sp.mammal.\*  
7: sp.mhc.\*  
8: sp.organelle.\*  
9: sp.phage.\*  
10: sp.plant.\*  
11: sp.rodent.\*  
12: sp.virus.\*  
13: sp.vertebrate.\*  
14: sp.unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	90.7	274	10	O82763 arabidopsis
2	37	86.0	239	12	O55741 chilo iride
3	36	83.7	284	10	Q39862 glycine max
4	36	83.7	319	10	Q40780 pimpinella
5	36	83.7	318	10	Q40781 pimpinella
6	36	83.7	302	10	Q40782 pimpinella
7	36	83.7	171	10	Q43441 glycine max
8	34	79.1	189	2	O67915 aquifex aeo
9	33	76.7	236	10	Q39927 helianthus
10	33	76.7	286	10	Q42437 lycopersico
11	32	74.4	972	3	O42975 schizosacch
12	32	74.4	2713	4	Q14660 homo sapien
13	32	74.4	2695	4	Q14643 homo sapien
14	32	74.4	1157	4	O95104 homo sapien
15	32	74.4	338	5	O17928 caenorhabdi
16	32	74.4	495	10	O22405 petroselinu
17	32	74.4	235	11	O63626 rattus norv
18	31	72.1	388	2	O54685 O54685
19	31	72.1	383	2	O48621 lactococcus
20	31	72.1	1698	5	O94438 chironomus
21	31	72.1	492	5	O18411 drosophila
22	31	72.1	583	5	O27472 caenorhabdi
23	31	72.1	355	5	O17930 caenorhabdi
24	31	72.1	771	5	O18174 caenorhabdi
25	31	72.1	401	5	O45907 caenorhabdi
26	31	72.1	2539	5	O96157 plasmodium
27	31	72.1	348	12	O41945 murine herp
28	31	72.1	422	12	O9YVW4 melanoplus
29	31	72.1	1100	13	O57576 cynops pyrr

30	31	72.1	846	13	O57577
31	31	72.1	300	13	Q9YGD9 oncothynchu
32	30	69.8	896	1	O58052 pyrococcus
33	30	69.8	128	1	O58297 pyrococcus
34	30	69.8	75	2	P94698 desulfovibr
35	30	69.8	92	2	Q9ZMC9 helicobacte
36	30	69.8	301	3	Q06701 saccharomyc
37	30	69.8	425	3	O14099 schizosacch
38	30	69.8	301	3	Q02799 saccharomyc
39	30	69.8	538	3	Q12582 candida mal
40	30	69.8	265	3	O74841 schizosacch
41	30	69.8	583	3	O14117 schizosacch
42	30	69.8	762	4	Q14965 homo sapien
43	30	69.8	436	4	Q14988 homo sapien
44	30	69.8	45	4	Q92714 homo sapien
45	30	69.8	1539	4	O15078 homo sapien

#### ALIGNMENTS

RESULT 1  
O82763 PRELIMINARY; PRT; 274 AA.  
ID AC O82763;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
DE HOMEOBOX PROTEIN, HAY9.  
GN T30L20.6 OR T20K9.1.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,  
RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;  
RA "Arabidopsis thaliana chromosome II BAC T30L20 genomic sequence";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,  
RA SYKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,  
RA SOMERVILLE C.R., VENTER J.C.;  
RL "Arabidopsis thaliana chromosome II BAC T20K9 genomic sequence";  
RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL; AC005617; AAC63591.1; -;  
DR EMBL; AC004786; AAC32427.1; -;  
DR PFAM; PF00046; homeobox; 1.  
DR PROSITE; PS00027; HOMEOBOX\_1; 1.  
KW Homeobox; DNA-binding; Nuclear protein.  
SQ SEQUENCE 274 AA; 29877 MW; 0797F7C1 CRC32;

Query Match 90.7%; Score 39; DB 10; Length 274;  
Best Local Similarity 85.7%; Pred. No. 2.3;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKKC 7

Db 174 CEFLKKC 180

RESULT 2

O55741 PRELIMINARY; PRT; 239 AA.  
ID O55741;  
AC O55741;

DT 01-JUN-1998 (Tremblrel. 06, Created)



CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M32491; G172177; -.  
DR EMBL; X79489; G496664; -.  
DR EMBL; Z35866; G536178; -.  
DR PIR; S45390; S45390.  
DR SGD; L0001446; PKC1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS00479; DAG\_PE\_BINDING\_DOMAIN; 2.  
DR PROSITE; PS50004; C2\_DOMAIN\_2; FALSE\_NEG.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PFAM; PF00069; pkinase; 1.  
DR PFAM; PF00130; DAG\_PE-bind; 2.  
DR PFAM; PF00168; C2; 1.  
DR PFAM; PF00433; pkinase\_C; 1.  
DR HSP; P04049; 1FAQ.  
KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING; ZINC;  
KW PHORBOL-ESTER BINDING; CELL CYCLE; DUPLICATION; CALCIUM-BINDING.  
FT DOMAIN 196 288  
FT DOMAIN 415 461  
FT DOMAIN 482 531  
FT DOMAIN 824 1083  
FT NP\_BIND 830 838  
FT BINDING 853 853  
FT ACT\_SITE 949 949  
FT VARIANT 958 958  
FT VARIANT 1023 1023  
FT CONFLICT 81 81  
FT CONFLICT 244 244  
FT CONFLICT 606 606  
FT CONFLICT 623 623  
FT CONFLICT 789 789  
SQ SEQUENCE 1151 AA; 131518 MW; 00E6AC9F CRC32;

Query Match 74.4%; Score 32; DB 1; Length 1151;  
Best Local Similarity 77.8%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Oy 1 CKFL--KKC 7  
Db 445 CKFLCHKKC 453  
|||||  
|||

RESULT 15  
PE38\_NPVOP  
ID PE38\_NPVOP STANDARD; PRT; 307 AA.  
AC P32512;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MAJOR IMMEDIATE EARLY PROTEIN (PE-38).  
GN PE38.  
OS ORGYIA PSEUDOTSUGATA MULTICAPSID POLYHEDROSIS VIRUS (OPMNPV).  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; BACULOVIRIDAE;  
OC NUCLEOPOLYHEDROVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92142537.  
RA THEILMANN D.A., STEWART S.;

RT "Tandemly repeated sequence at the 3' end of the IE-2 gene of the  
RT baculovirus Orgyia pseudotsugata multcapsid nuclear polyhedrosis  
RT virus is an enhancer element.";  
RL VIROLOGY 187:97-106(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92711300.  
RA ARENS C.H., RUSSELL R.R., FUNK C.J., EVANS J., HARWOOD S.,  
RA ROHRMANN G.F.;  
RT "The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear  
RT polyhedrosis virus genome.";  
RL VIROLOGY 229:381-399(1997).  
CC -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M83827; G332542; -.  
DR EMBL; U75930; G1911398; -.  
DR PIR; A42191; WMNV38.  
DR PROSITE; PS00518; ZINC\_FINGER\_C3HC4; 1.  
DR PFAM; PF00097; zf-C3HC4; 1.  
KW EARLY PROTEIN; ZINC-FINGER; DNA-BINDING.  
FT ZN\_FING 39 91  
SQ SEQUENCE 307 AA; 34714 MW; E712CE99 CRC32;

Query Match 74.4%; Score 32; DB 1; Length 307;  
Best Local Similarity 83.3%; Pred. No. 30;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CKFLKK 6  
Db 109 CKFIKK 114  
|||||  
|||

Search completed: September 7, 1999, 23:59:11  
Job time: 523 sec

```

DR EMBL: J05510; G204675; -
DR EMBL: J05510; G204674; -
DR EMBL: M64699; G204967; -
DR EMBL: M64698; G204969; -
DR PIR: A36579; A36579.
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; PHOSPHORYLATION;
KW ENDOPLASMIC RETICULUM; IONIC CHANNEL; ION TRANSPORT; CALCIUM CHANNEL;
KW ALTERNATIVE SPLICING.
FT DOMAIN 1 2273 CYTOPLASMIC (POTENTIAL).
FT TRANSNM 2274 2294 M1 (POTENTIAL).
FT TRANSNM 2308 2326 M2 (POTENTIAL).
FT TRANSNM 2334 2356 M3 (POTENTIAL).
FT TRANSNM 2365 2387 M4 (POTENTIAL).
FT TRANSNM 2391 2407 M5 (POTENTIAL).
FT TRANSNM 2440 2462 M6 (POTENTIAL).
FT TRANSNM 2530 2549 M7 (POTENTIAL).
FT TRANSNM 2570 2589 M8 (POTENTIAL).
FT DOMAIN 2590 2749 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 1589 1589 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT MOD_RES 1755 1755 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT VARIANT 1372 1372 MISSING (IN ALL, BUT P17 CLONES).
FT VARSPLIC 322 336 MISSING (IN VARIANT).
FT VARSPLIC 1693 1731 MISSING (IN SHORT FORM).
FT CONFLICT 1716 1716 E -> EQ (IN REF. 2).
FT CONFLICT 1763 1763 P -> R (IN REF. 2).
SQ SEQUENCE 2749 AA; 313132 MW; A64DE4B1 CRC32;

Query Match 74.4%; Score 32; DB 1; Length 2749;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKCC 7
DB 1321 KFIKCC 1326

RESULT 13
KPC1_CANAL STANDARD; PRT; 1097 AA.
ID KPC1_CANAL
AC P43057;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE PROTEIN KINASE C-LIKE 1 (EC 2.7.1.-) (PKC 1).
GN PKC1.
OS CANDIDA ALBICANS (YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC CANDIDACEAE; CANDIDA.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 10261;
RX MEDLINE: 96408767.
RA PARAVICINI G., MENDOZA A., ANTONSSON B., COOPER M., LOSBERGER C.,
RA PAYTON M.A.;
RT "The Candida albicans PKC1 gene encodes a protein kinase C homolog
RT necessary for cellular integrity but not dimorphism.";
RL YEAST 12:741-756(1996).
CC -!- FUNCTION: NECESSARY FOR OSMOTIC STABILITY.
CC -!- SIMILARITY: CONTAINS TWO COPIES OF THE ZINC-DEPENDENT PHORBOL-
CC ESTER AND DAG BINDING DOMAIN.
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES. BELONGS TO THE PKC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X81142; G832908; -

```

---

```

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00479; DAG_PE_BINDING_DOMAIN; 2.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00069; pkinase; 1.
DR PFAM; PF00130; DAG_PE-bind; 2.
DR PFAM; PF00433; pkinase_C; 1.
DR HSSP; Q63450; 1A06.
KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING; ZINC;
KW PHORBOL-ESTER BINDING; DUPLICATION; PHORBOL-ESTER AND DAG BINDING (BY
FT DOMAIN 416 462 SIMILARITY).
FT DOMAIN 477 530 PHORBOL-ESTER AND DAG BINDING (BY
FT DOMAIN 770 1029 SIMILARITY).
FT NP_BIND 776 784 PROTEIN KINASE.
FT BINDING 799 799 ATP (BY SIMILARITY).
FT ACT_SITE 895 895 BY SIMILARITY.
SQ SEQUENCE 1097 AA; 125312 MW; 9DC0A27E CRC32;

Query Match 74.4%; Score 32; DB 1; Length 1097;
Best Local Similarity 77.8%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 CKFL--KKC 7
DB 446 CKFLCHKKC 454

RESULT 14
KPC1_YEAST STANDARD; PRT; 1151 AA.
ID KPC1_YEAST
AC P24583;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROTEIN KINASE C-LIKE 1 (EC 2.7.1.-) (PKC 1).
GN PKC1 OR STT1 OR HPO2 OR YBL105C OR YBL0807.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=EG123;
RX MEDLINE: 90322419.
RA LEVIN D.E., FIELDS F.O., KUNISAWA R., BISHOP J.M., THORNER J.;
RT "A candidate protein kinase C gene, PKC1, is required for the S.
RT cerevisiae cell cycle.";
RL CELL 62:213-224(1990).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE: 96076635.
RA OBERMAIER B., GASSENHUBER J., PIRAVANDI E., DOMDEY H.;
RT "Sequence analysis of a 78.6 kb segment of the left end of
RT Saccharomyces cerevisiae chromosome II.";
RL YEAST 11:1103-1112(1995).
[3]
RN VARIANTS CLY5 AND CLY7.
RP MEDLINE: 97279226.
RX BAYMILLER J., MCCULLOUGH J.E.;
RA "Saccharomyces cerevisiae cell lysis mutations cly5 and cly7 define
RT temperature-sensitive alleles of PKC1, the gene encoding yeast
RT protein kinase C.";
RL YEAST 13:305-312(1997).
CC -!- FUNCTION: REQUIRED FOR CELL GROWTH AND FOR THE G2->M TRANSITION OF
CC THE CELL DIVISION CYCLE. MEDIATES A PROTEIN KINASE CASCADE; IT
CC ACTIVATES BCK1 WHICH ITSELF ACTIVATES MKK1/MKK2.
CC -!- SIMILARITY: CONTAINS TWO COPIES OF THE ZINC-DEPENDENT PHORBOL-
CC ESTER AND DAG BINDING DOMAIN.
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES. BELONGS TO THE PKC FAMILY.

```

RC TISSUE-PURKINJE CELLS;  
 RX MEDLINE; 90044039.  
 RA FURUICHI T., YOSHIKAWA S., MIYAWAKI A., WADA K., MAEDA N.,  
 RA MIKOSHIBA K.;  
 RT "Primary structure and functional expression of the inositol 1,4,5-  
 trisphosphate-binding protein P400.";  
 RL NATURE 342:32-38(1989).  
 RN (2)  
 RC SEQUENCE FROM N.A.  
 RP STRAIN-ICR; TISSUE-CEREBELLUM;  
 RX MEDLINE; 89345101.  
 RA FURUICHI T., YOSHIKAWA S., MIKOSHIBA K.;  
 RT "Nucleotide sequence of cDNA encoding P400 protein in the mouse  
 cerebellum.";  
 RL NUCLEIC ACIDS RES. 17:5385-5386(1989).  
 RN (3)  
 RP SEQUENCE OF 318-332 AND 1692-1731 FROM N.A., AND ALTERNATIVE SPLICING.  
 RC STRAIN-JCR;  
 RX MEDLINE; 91296797.  
 RA NAKAGAWA T., OKANO H., FURUICHI T., ARUGA J., MIKOSHIBA K.;  
 RT "The subtypes of the mouse inositol 1,4,5-trisphosphate receptor are  
 expressed in a tissue-specific and developmentally specific manner.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:6244-6248(1991).  
 CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND  
 CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
 CC THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL  
 CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-  
 CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE  
 CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
 CC RETICULUM.  
 CC -1- PTM: PHOSPHORYLATED BY CYCLIC-AMP KINASE. PHOSPHORYLATION PREVENTS  
 CC THE LIGAND-INDUCED OPENING OF THE CALCIUM CHANNELS.  
 CC -1- CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE RECEPTOR, MOST  
 CC PROBABLY BY INTERACTING WITH A DISTINCT CALCIUM-BINDING PROTEIN  
 CC WHICH THEN INHIBITS THE RECEPTOR.  
 CC -1- ALTERNATIVE PRODUCTS: ADDITIONAL SUBTYPES OF INSP3R ARISE BY  
 CC ALTERNATIVE SPLICING OF THE SAME GENE.  
 CC -1- SIMILARITY: TO RYANODINE RECEPTOR.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X15373; G53569; .  
 DR EMBL; M75986; G198412; .  
 DR EMBL; M75987; G554171; .  
 DR PIR; S04844; ACMSIT.  
 DR MGD; MGI:96623; ITPR1.  
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; PHOSPHORYLATION;  
 KW ENDOPLASMIC RETICULUM; IONIC CHANNEL; ION TRANSPORT; CALCIUM CHANNEL;  
 KW ALTERNATIVE SPLICING.  
 FT DOMAIN 1 2273  
 FT TRANSMEM 2274 2294  
 FT TRANSMEM 2308 2326  
 FT TRANSMEM 2334 2356  
 FT TRANSMEM 2365 2387  
 FT TRANSMEM 2391 2407  
 FT TRANSMEM 2440 2462  
 FT TRANSMEM 2530 2549  
 FT TRANSMEM 2570 2589  
 FT DOMAIN 2590 2749  
 FT VARSPIC 318 332  
 FT VARSPIC 1692 1731  
 FT VARSPIC 1715 1715  
 FT VARSPIC 1715 1715  
 FT VARSPIC 1715 1715  
 FT MOD\_RES 1588 1588  
 FT MOD\_RES 1755 1755

SQL SEQUENCE 2749 AA; 313193 MW; B19B78F2 CRC32;  
 Query Match 74.4%; Score 32; DB 1; Length 2749;  
 Best Local Similarity 83.3%; Pred. No. 2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KFLKKC 7  
 DB 1321 KFIKKC 1326  
 RESULT 12  
 IP3R\_RAT  
 ID IP3R\_RAT STANDARD; PRT; 2749 AA.  
 AC P29994;  
 DT 01-APR-1993 (REL. 25, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN TYPE 1 RECEPTOR  
 DE (INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN P(400)) (TYPE 1 INSP3  
 DE RECEPTOR).  
 GN INSP3R.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90324264.  
 RA MIGNERY G.A., NEWTON C.L., ARCHER B.T. III, SUEHDHOF T.C.;  
 RT "Structure and expression of the rat inositol 1,4,5-trisphosphate  
 RT receptor.";  
 RL J. BIOL. CHEM. 265:12679-12685(1990).  
 RN [2]  
 RP SEQUENCE OF 1652-1824 FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE-BRAIN;  
 RX MEDLINE; 91187909.  
 RA DULOFF S.K., FERRIS C.D., DONATH C., FISCHER G.A., MUNEMITSU S.,  
 RA ULLRICH A., SNYDER S.H., ROSS C.A.;  
 RT "Inositol 1,4,5-trisphosphate receptors: distinct neuronal and  
 RT nonneuronal forms derived by alternative splicing differ in  
 RT phosphorylation.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:2951-2955(1991).  
 CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND  
 CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
 CC THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL  
 CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-  
 CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE  
 CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
 CC RETICULUM.  
 CC -1- TISSUE SPECIFICITY: THE LONGER ISOFORM IS FOUND IN THE BRAIN WHILE  
 CC THE SHORTER ISOFORM IS FOUND IN THE FETAL BRAIN AND PERIPHERAL  
 CC TISSUES.  
 CC -1- PTM: PHOSPHORYLATED BY CYCLIC-AMP KINASE. PHOSPHORYLATION PREVENTS  
 CC THE LIGAND-INDUCED OPENING OF THE CALCIUM CHANNELS.  
 CC -1- CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE RECEPTOR, MOST  
 CC PROBABLY BY INTERACTING WITH A DISTINCT CALCIUM-BINDING PROTEIN  
 CC WHICH THEN INHIBITS THE RECEPTOR.  
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS, A LONG (NEURONAL) AND A SHORT  
 CC (NONNEURONAL) FORM, ARE PRODUCED BY ALTERNATIVE SPLICING OF THE  
 CC SAME GENE.  
 CC -1- SIMILARITY: TO RYANODINE RECEPTOR.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL: X68145; G16180; -  
CC DR EMBL: X68146; G16332; -  
CC DR EMBL: Z19602; G22759; -  
CC DR EMBL: M90394; G166752; -  
CC DR EMBL: Z87342; E327013; -  
CC DR PIR: S28566; S28566;  
CC DR PIR: S31424; S31424;  
CC DR PROSITE: PS00027; HOMEBOX\_1; 1.  
CC DR PROSITE: PS00071; HOMEBOX\_2; 1.  
CC DR PFAM: PF00046; homeobox; 1.  
CC DR HSP: P01366; IAKH.  
CC DR TRANSFAC: T01475; -  
CC DR TRANSFAC: T02051; -  
CC DR HOMEBOX: DNA-BINDING; NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN.  
CC KW HOMEBOX: 126 185 HOMEBOX.  
CC FT DOMAIN 193 214 LEUCINE-ZIPPER.  
CC FT DOMAIN 216 216 A -> V (IN REF. 2 AND 3).  
CC FT CONFLICT 216 216 A -> V (IN REF. 2 AND 3).  
CC SQ SEQUENCE 284 AA; 31852 MW; E0EDC28 CRC32;

Query Match 76.7%; Score 33; DB 1; Length 284;  
Best Local Similarity 57.1%; Pred. No. 18;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKCC 7  
|:|:|:  
DB 190 CEFLRRC 196

RESULT 9  
HT14\_ARATH  
ID HT14\_ARATH STANDARD; PRT; 165 AA.  
AC P46665;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE HOMEBOX-LEUCINE ZIPPER PROTEIN HT14 (HD-ZIP PROTEIN 14) (FRAGMENT).  
GN HT14.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RX MEDLINE; 94359937.  
RA SCHENA M., DAVIS R.W.;  
RT "Structure of homeobox-leucine zipper genes suggests a model for the  
RT evolution of gene families";  
RL PROC. NATL. ACAD. SCI. U.S.A. 91:8393-8397(1994).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
CC -!- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL: U09334; G54985;  
CC DR PROSITE: PS00027; HOMEBOX\_1; 1.  
CC DR PROSITE: PS00071; HOMEBOX\_2; 1.  
CC DR PFAM: PF00046; homeobox; 1.  
CC DR HSP: P01366; LYRN.

KW HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN.  
FT NON\_TER 1 1  
FT DNA\_BIND 16 75 HOMEBOX.  
FT DOMAIN 83 104 LEUCINE-ZIPPER.  
SQ SEQUENCE 165 AA; 18742 MW; C1A33482 CRC32;

Query Match 76.7%; Score 33; DB 1; Length 165;  
Best Local Similarity 57.1%; Pred. No. 11;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKCC 7  
|:|:|:  
DB 80 CEYLKRC 86

RESULT 10  
BRL\_RANBP  
ID BRL\_RANBP STANDARD; PRT; 24 AA.  
AC P32423;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE BREVININ-1.  
OS RANA BREVIPODA PORSA (FROG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
OC NEOBATRACHIA; RANOIDEA; RANIDAE; RANINAE; RANA.  
RN [1]  
RP SEQUENCE.  
RC TISSUE-SKIN;  
RX MEDLINE; 93080555.  
RA MORIKAWA N., HAGIWARA K., NAKAJIMA T.;  
RT "Brevinin-1 and -2, unique antimicrobial peptides from the skin of  
RT the frog, Rana brevipoada porsa";  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 189:184-190(1992).  
CC -!- FUNCTION: SHOWS ANTIBACTERIAL ACTIVITY AGAINST REPRESENTATIVE  
CC GRAM-NEGATIVE AND GRAM-POSITIVE BACTERIAL SPECIES, AND A VERY  
CC HIGH HEMOLYTIC ACTIVITY.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- TISSUE SPECIFICITY: SKIN.  
CC -!- SIMILARITY: BELONGS THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN  
CC FAMILY.  
CC PIR: JCL355; JCL355.  
CC AMPHIBIAN SKIN; ANTIBIOTIC; HEMOLYSIS.  
FT DISULFID 18 24 BY SIMILARITY.  
FT SEQUENCE 24 AA; 2531 MW; 264D7A60 CRC32;

Query Match 74.4%; Score 32; DB 1; Length 24;  
Best Local Similarity 71.4%; Pred. No. 3.2;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFLKCC 7  
|:|:|:  
DB 18 CKITKCC 24

RESULT 11  
IP3R\_MOUSE  
ID IP3R\_MOUSE STANDARD; PRT; 2749 AA.  
AC P11881;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN TYPE 1 RECEPTOR  
DE (INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN P(400)) (TYPE 1 INSP3  
DE RECEPTOR).  
GN INSP3R OR ITPRI.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.



FT CONFLICT 950 950 V -> L (IN REF. 1).  
FT CONFLICT 1005 1005 S -> G (IN REF. 1).  
SQ SEQUENCE 1089 AA; 122727 MW; 82FAE940 CRC32;

Query Match 79.1%; Score 34; DB 1; Length 1089;

Best Local Similarity 71.4%; Pred. No. 39;

Matches 5; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

OY 1 CKFLKCC 7

I::|::|

Db 450 CKHIKC 456

RESULT 4

ATH4 ARATH

ID ATH4 ARATH STANDARD; PRT; 318 AA.

AC P92953; (REL. 36, CREATED)

DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)

DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE HOMEOBOX-LEUCINE ZIPPER PROTEIN ATHB-4 (HD-ZIP PROTEIN ATHB-4).

GN ATHB-4 OR TL3E15.8.

OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).

OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; TRACHEOPHYTA;

OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;

OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.

RN [1]

RP SEQUENCE FROM N.A.

RA STEINDLER C., MORELLI G., RUBERTI I.;

RT "Nucleotide sequence of the Arabidopsis ATHB-4 gene encoding an HD-Zip

protein related to ATHB-2.";

RL (IN) PLANT GENE REGISTER PGR97-021.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;

RA ROUNSELEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,

RA RYKES S.M., MASON T.M., KERLAVAGE A.R., ADAMS M.D., SOMERVILLE C.R.,

RA VENTER J.C.;

RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).

CC -1- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; Y09582; E283562; ..

DR EMBL; AC002388; G2344893; ..

DR PROSITE; PS00027; HOMEOBOX\_1; 1.

DR PROSITE; PS50071; HOMEOBOX\_2; 1.

DR PFAM; PF00046; homeobox; 1.

DR HSSP; P01366; JAKH

KW HOMEOBOX; DNA-BINDING; NUCLEAR PROTEIN.

FT DOMAIN 34 39 POLY-SER.

FT DNA\_BIND 160 219 HOMEOBOX.

FT DOMAIN 289 294 POLY-THR.

SQ SEQUENCE 318 AA; 34816 MW; F4522E0 CRC32;

Query Match 76.7%; Score 33; DB 1; Length 318;

Best Local Similarity 57.1%; Pred. No. 20;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 CKFLKCC 7

I::|::|

Db 224 CEYLKRC 230

RESULT 5

HAT1 ARATH

ID HAT1 ARATH STANDARD; PRT; 282 AA.

AC P46600; (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DE HOMEOBOX-LEUCINE ZIPPER PROTEIN HAT1 (HD-ZIP PROTEIN 1).

GN HAT1.

OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).

OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMERYOPHYTA; TRACHEOPHYTA;

OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;

OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;

RX MEDLINE; 94359937.

RA SCHENA M., DAVIS R.W.;

RT "Structure of homeobox-leucine zipper genes suggests a model for the

evolution of gene families.";

RL PROC. NATL. ACAD. SCI. U.S.A. 91:8393-8397(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;

RA BEVAN M., STIEKEMA W., MURPHY G., WAMBUIT R., POHL T., TERRY N.,

RA KREIS M., KAVANAGH T., ENTIAN K.-D., RIEGER M., JAMES R.,

RA PUIGDOMENECH P., HATZOPOULOS P., OBERMAIER B., DUESTERHOFT A.,

RA JONES J., PALME K., ANSORGE W., DELSENY M., BANCROFT I., MEWES H.-W.,

RA SCHUELLER C., CHALWATZIS N.;

RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).

CC -1- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; U09332; G549883; ..

DR EMBL; U09333; G549884; ..

DR EMBL; Z97343; E327063; ..

DR PROSITE; PS00027; HOMEOBOX\_1; 1.

DR PROSITE; PS50071; HOMEOBOX\_2; 1.

DR PFAM; PF00046; homeobox; 1.

DR HSSP; P01366; LYRN.

KW HOMEOBOX; DNA-BINDING; NUCLEAR PROTEIN.

FT DNA\_BIND 132 191 HOMEOBOX.

FT DOMAIN 199 220 LEUCINE-ZIPPER.

SQ SEQUENCE 282 AA; 31554 MW; 2C7C499C CRC32;

Query Match 76.7%; Score 33; DB 1; Length 282;

Best Local Similarity 57.1%; Pred. No. 18;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 CKFLKCC 7

I::|::|

Db 196 CEYLKRC 202

RESULT 6

HAT2 ARATH

ID HAT2 ARATH STANDARD; PRT; 208 AA.

AC P46601; (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DE HOMEOBOX-LEUCINE ZIPPER PROTEIN HAT2 (HD-ZIP PROTEIN 2) (FRAGMENT).

GN HAT2.

OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).

CC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 CC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 CC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=CV. COLUMBIA;  
 RX MEDLINE; 94359937.  
 RA SCHENA M., DAVIS R.W.;  
 RT "Structure of homeobox-leucine zipper genes suggests a model for the  
 evolution of gene families.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:8393-8397(1994).  
 [2]  
 RN SEQUENCE OF 166-221 FROM N.A.  
 RP STRAIN=CV. COLUMBIA;  
 RX MEDLINE; 92237275.  
 RA SCHENA M., DAVIS R.W.;  
 RT "HD-Zip proteins: members of an Arabidopsis homeodomain protein  
 superfamily.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:3894-3898(1992).  
 [3]  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CC -1- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U09336; G549887; -;  
 DR EMBL; U09337; G549888; -;  
 DR EMBL; M90417; G166756; -;  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 DR PFAM; PF00046; Homeobox; 1.  
 DR HSP; P01366; LYRN.  
 KW HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN.  
 FT DNA\_BIND 123 182 HOMEBOX.  
 FT DOMAIN 190 211 LEUCINE-ZIPPER.  
 FT CONFLICT 221 221 Y -> S (IN REF. 2).  
 SQ SEQUENCE 278 AA; 30729 MW; D50E698D CRC32;  
 -----  
 Query Match 90.7%; Score 39; DB 1; Length 278;  
 Best Local Similarity 85.7%; Pred. No. 1.4;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CKFKKC 7  
 Db 187 CEFLKC 193  
 -----  
 RESULT 3  
 PGDS\_MOUSE STANDARD; PRT; 1089 AA.  
 AC P26618; O62046;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE ALPHA PLATELET-DERIVED GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112)  
 DE (PDGF-R-ALPHA).  
 GN PDGFRA.  
 OS MUS MUSCULUS (MOUSE).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STILES C.D., WANG C.;  
 RX MEDLINE; 91061789.  
 RT "Retinoic acid promotes transcription of the platelet-derived growth  
 factor alpha-receptor gene.";  
 RL MOL. CELL. BIOL. 10:6781-6784(1990).

RN SEQUENCE FROM N.A.  
 RX MEDLINE; 92334866.  
 RA DO M.S., FITZER-ATTAS C., GUBBAY J., GREENFELD L., FELDMAN M.,  
 RA EISENBACH L.;  
 RT "Mouse platelet-derived growth factor alpha receptor: sequence,  
 RT tissue-specific expression and correlation with metastatic  
 RT phenotype.";  
 RL ONCOGENE 7:1567-1575(1992).  
 CC -1- FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND  
 CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR CAN BIND  
 CC EITHER PDGF-A OR PDGF-B.  
 CC -1- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA  
 CC SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. THE  
 CC EXTRACELLULAR DOMAIN CONTAINS FIVE IG-FOLD DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M84607; G200275; -;  
 DR EMBL; M57683; G199784; -;  
 DR PIR; S33727; S33727.  
 DR MGD; MGI:97530; PDGFRA.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PFAM; PF00047; Ig; 2.  
 DR HSP; P11362; IAGW.  
 KW TYROSINE-PROTEIN KINASE; RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW TRANSFERASE; PHOSPHORYLATION; ATP-BINDING; IMMUNOGLOBULIN FOLD;  
 KW SIGNAL.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 1089 ALPHA PLATELET-DERIVED GROWTH FACTOR  
 FT RECEPTOR.  
 FT DOMAIN 25 525 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 526 549 POTENTIAL.  
 FT DOMAIN 550 1089 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 599 607 PROTEIN KINASE.  
 FT BINDING 627 627 ATP (BY SIMILARITY).  
 FT ACT\_SITE 818 818 ATP (BY SIMILARITY).  
 FT CARBOHYD 42 42 BY SIMILARITY.  
 FT CARBOHYD 76 76 POTENTIAL.  
 FT CARBOHYD 89 89 POTENTIAL.  
 FT CARBOHYD 103 103 POTENTIAL.  
 FT CARBOHYD 179 179 POTENTIAL.  
 FT CARBOHYD 353 353 POTENTIAL.  
 FT CARBOHYD 359 359 POTENTIAL.  
 FT CARBOHYD 458 458 POTENTIAL.  
 FT CARBOHYD 468 468 POTENTIAL.  
 FT CARBOHYD 506 506 POTENTIAL.  
 FT CONFLICT 192 192 T -> A (IN REF. 1).  
 FT CONFLICT 202 202 E -> A (IN REF. 1).  
 FT CONFLICT 252 252 E -> G (IN REF. 1).  
 FT CONFLICT 271 271 L -> V (IN REF. 1).  
 FT CONFLICT 322 322 G -> S (IN REF. 1).  
 FT CONFLICT 326 326 A -> P (IN REF. 1).  
 FT CONFLICT 439 440 GT -> EG (IN REF. 1).  
 FT CONFLICT 529 529 A -> E (IN REF. 1).  
 FT CONFLICT 737 737 A -> D (IN REF. 1).  
 FT CONFLICT 849 849 Y -> D (IN REF. 1).  
 FT CONFLICT 936 936 E -> D (IN REF. 1).

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:10 ; Search time 71.09 Seconds  
(without alignments)  
2.783 Million cell updates/sec

Title: US-09-124-280a-30

Perfect score: 43

Sequence: 1 CKFLKKC 7

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	90.7	274	1 HAT9_ARATH	P46603 arabidopsis
2	39	90.7	278	1 HT22_ARATH	P46604 arabidopsis
3	34	79.1	1089	1 PGDS_MOUSE	P26618 mus musculus
4	33	76.7	318	1 ATH4_ARATH	P92953 arabidopsis
5	33	76.7	282	1 HAT1_ARATH	P46600 arabidopsis
6	33	76.7	208	1 HAT2_ARATH	P46601 arabidopsis
7	33	76.7	315	1 HAT3_ARATH	P46602 arabidopsis
8	33	76.7	284	1 HAT4_ARATH	Q05466 arabidopsis
9	33	76.7	165	1 HT14_ARATH	P46665 arabidopsis
10	32	74.4	24	1 BRL_FANBP	P32423 rana brevip
11	32	74.4	2749	1 IP3R_MOUSE	P11881 mus musculus
12	32	74.4	2749	1 IP3R_RAT	P29994 rattus norv
13	32	74.4	1097	1 KPCI_CANAL	P43057 candida alb
14	32	74.4	1151	1 KPCI_YEAST	P24583 saccharomyc
15	32	74.4	307	1 PE38_NPVOP	P32512 oryza pseu
16	32	74.4	1089	1 PGDS_HUMAN	P16234 homo sapien
17	32	74.4	1088	1 PGDS_RAT	P20786 rattus norv
18	32	74.4	223	1 YZL2_EBV	P03205 Epstein-bar
19	31	72.1	244	1 AEG1_MOUSE	Q03401 mus musculu
20	31	72.1	1700	1 BAR3_CHITE	Q03376 chironomus
21	31	72.1	487	1 HEP_DROME	Q23977 drosophila
22	31	72.1	853	1 NCAL_BOVIN	P31836 bos taurus
23	31	72.1	1091	1 NCAL_CHICK	P13590 gallus gall
24	31	72.1	848	1 NCAL_HUMAN	P13591 homo sapien
25	31	72.1	1115	1 NCAL_MOUSE	P13595 mus musculu
26	31	72.1	858	1 NCAL_RAT	P13596 rattus norv
27	31	72.1	403	1 NOF_DROME	P15297 drosophila
28	31	72.1	984	1 NOF_DROME	P16320 drosophila
29	31	72.1	197	1 VS11_ROT5	P19715 porcine rot
30	31	72.1	816	1 YG3A_YEAST	P53278 saccharomyc
31	31	72.1	376	1 YOT3_CAEEL	Q09315 caenorhabdi
32	30	69.8	538	1 CP5D_CANNA	P16141 candida mal
33	30	69.8	49	1 DIS1_ECHCA	P17347 echis carin
34	30	69.8	4092	1 DYHC_YEAST	P36022 saccharomyc
35	30	69.8	1147	1 MYSB_ACACA	P19706 acanthameb
36	30	69.8	1088	1 NCAL_XENLA	P16170 xenopus lae
37	30	69.8	1092	1 NCA2_XENLA	P36335 xenopus lae
38	30	69.8	1059	1 POL2_DROME	P20825 drosophila
39	30	69.8	285	1 RL2_MYCPE	P47400 mycoplasma
40	30	69.8	287	1 RL2_MYCPE	P75577 mycoplasma
41	30	69.8	1871	1 SEX_HUMAN	P51805 homo sapien
42	30	69.8	169	1 VG12_HSV5A	P24915 herpesvirus
43	30	69.8	105	1 YA83_HAEIN	Q57407 haemophilus

44 29 67.4 71 1 BRL\_FANBP P32412 rana esculi  
45 29 67.4 370 1 COMT\_CLABR O23760 clarkia bre

#### ALIGNMENTS

```

RESULT 1
HAT9_ARATH
ID HAT9_ARATH STANDARD; PRT; 274 AA.
AC P46603;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HOMOBOX-LEUCINE ZIPPER PROTEIN HAT9 (HD-ZIP PROTEIN 9).
GN HAT9.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLIPHTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE; 94359937.
RA SCHENA M., DAVIS R.W.;
RT "Structure of homeobox-leucine zipper genes suggests a model for the
   evolution of gene families."
RL PROC. NATL. ACAD. SCI. U.S.A. 91:8393-8397(1994).
CC 1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC 1- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U09341; G549892; -
CC DR EMBL; U09342; G549893; -
CC DR PROSITE; PS00027; HOMOBOX_1; 1.
CC DR PROSITE; PS00071; HOMOBOX_2; 1.
CC DR PFAM; PF00046; homeobox; 1.
CC DR HSSP; P01366; LAKH.
CC DR HOMOBOX; DNA-BINDING; NUCLEAR PROTEIN.
CC FT DNA_BIND 110 169 HOMOBOX.
CC FT DOMAIN 177 198 LEUCINE-ZIPPER.
CC FT DOMAIN 90 93 POLY-GLU.
CC FT DOMAIN 226 240 POLY-GLY.
CC FT CONFLICT 90 90 E -> V (IN G549893).
CC SQ SEQUENCE 274 AA; 29867 MW; DA3C3355 CRC32;

Query Match 90.7%; Score 39; DB 1; Length 274;
Best Local Similarity 85.7%; Pred. No. 1.4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKKC 7
DB 174 CEFLKKC 180

RESULT 2
HT22_ARATH
ID HT22_ARATH STANDARD; PRT; 278 AA.
AC P46604;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HOMOBOX-LEUCINE ZIPPER PROTEIN HAT22 (HD-ZIP PROTEIN 22).
GN HAT22.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).

```





A:Accession: A42191  
A:Molecule type: DNA  
A:Residues: 1-1151 <OBE>  
A:Cross-references: EMBL:X79489; NID:g496661; PID:g496664  
A:Experimental source: strain S288C  
R:Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E.  
C:Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus PE-38 protein  
C:Keywords: DNA binding; early protein

Query Match 74.4%; Score 32; DB 1; Length 307;  
Best Local Similarity 83.3%; Pred. No. 48;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKK 6  
|||||  
Db 109 CKFIKK 114

RESULT 14  
S47220  
protein kinase C (EC 2.7.1.1-) PKC1 - yeast (Candida albicans)  
C:Species: Candida albicans  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Mar-1998  
C:Accession: S47220; S71641  
R:Paravicini, G.; Mendoza, A.; Autoussou, B.; Cooper, M.; Losberger, C.; Payton, M.  
submitted to the EMBL Data Library, August 1994  
A:Description: The Candida albicans PKC1 gene encodes a protein kinase C (PKC) homolog  
A:Reference number: S47220  
A:Accession: S47220  
A:Molecule type: DNA  
A:Residues: 1-1097 <PAR>  
A:Cross-references: EMBL:X81142; NID:g535094; PID:g832908  
A:Experimental source: strain 10261  
R:Paravicini, G.; Mendoza, A.; Antonsson, B.; Cooper, M.; Losberger, C.; Payton, M.A.  
Yeast 12, 741-756, 1996  
A:Title: The Candida albicans PKC1 gene encodes a protein kinase C homolog necessary for  
A:Reference number: S71641; MUID:96408767  
A:Accession: S71641  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 396-531;764-1097 <PAW>  
A:Cross-references: EMBL:X81142  
C:Genetics:  
A:Gene: PKC1  
A:Superfamily: yeast protein kinase C; protein kinase C zinc-binding repeat homology; ph  
C:Keywords: ATP; duplication: phospholipid binding; phosphotransferase; serine/threonine  
F:416-462/Domain: protein kinase C zinc-binding repeat homology <K21>  
F:481-530/Domain: protein kinase C zinc-binding repeat homology <K22>  
F:768-1029/Domain: protein kinase C zinc-binding repeat homology <KIN>  
F:776-784/Region: protein kinase ATP-binding motif  
F:799,895/Active site: Lys, Asp #status predicted

Query Match 74.4%; Score 32; DB 2; Length 1097;  
Best Local Similarity 77.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 CKFL--KKC 7  
|||||  
Db 446 CKFLCHKKC 454

RESULT 15  
S45390  
protein kinase C (EC 2.7.1.1-) PKC1 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YBL0807; protein YBL105C  
C:Species: Saccharomyces cerevisiae  
C:Date: 09-Jun-1994 #sequence\_revision 09-Sep-1994 #text\_change 06-Feb-1998  
C:Accession: S45390; S45848; S12305; S59187  
R:Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.  
submitted to the EMBL Data Library, May 1994  
A:Description: Sequence analysis of a 78,6 kb segment of the left end of Saccharomyces ce  
A:Reference number: S45397  
A:Accession: S45390

A:Molecule type: DNA  
A:Residues: 1-1151 <DOM>  
A:Cross-references: EMBL:X35866; NID:g536177; PID:g536178; MIPS:YBL105C  
R:Levin, D.E.; Fields, F.O.; Kunisawa, R.; Bishop, J.M.; Thorne, J.  
Cell 62, 213-224, 1990  
A:Title: A candidate protein kinase C gene, PKC1, is required for the S. cerevisiae c  
A:Reference number: S12305; MUID:90322419  
A:Accession: S12305  
A:Molecule type: DNA  
A:Residues: 1-80,'C',82-243,'S',245-605,'E',607-622,'P',624-788,'A',790-1151 <LEV>  
A:Cross-references: EMBL:IM32491; NID:gl72176; PID:gl72177  
A:Experimental source: strain EG123  
R:Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.  
Yeast 11, 1103-1112, 1995  
A:Title: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cere  
A:Reference number: S59184; MUID:96076635  
A:Accession: S59187  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1151 <ORW>  
A:Cross-references: EMBL:X79489; NID:g496661; PID:g496664  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, May 1994  
C:Genetics:  
A:Gene: SGD:PKC1  
A:Cross-references: SGD:S0000201; MIPS:YBL105C  
A:Map position: 2L  
C:Function:  
A:Description: phosphotransferase; serine/threonine-specific protein kinase  
A:Note: cells depleted of PKC1 product arrest cell division at a point subsequent to  
C:Superfamily: yeast protein kinase C; protein kinase C zinc-binding repeat homology;  
C:Keywords: ATP; cell cycle control; phosphotransferase; serine/threonine-specific pr  
F:415-461/Domain: protein kinase C zinc-binding repeat homology <K21>  
F:482-531/Domain: protein kinase C zinc-binding repeat homology <K22>  
F:822-1083/Domain: protein kinase homology <KIN>  
F:830-838/Region: protein kinase ATP-binding motif  
F:853,949/Active site: Lys, Asp #status predicted

Query Match 74.4%; Score 32; DB 2; Length 1151;  
Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 CKFL--KKC 7  
|||||  
Db 445 CKFLCHKKC 453

Search completed: September 7, 1999, 23:22:45  
Job time: 952 sec

A:Note: in the authors' translation an additional residue, Val, is shown after position R;Herrén, B.; Weyer, K.A.; Rouge, M.; Loetscher, P.; Pech, M.  
 Biochim. Biophys. Acta 1173, 294-302, 1993  
 A:Title: Conservation in sequence and affinity of human and rodent PDGF ligands and receptors  
 A:Reference number: S33764; MUID:93305723  
 A:Accession: S33767  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 33-149, 'R', 151-518, 'T', 520-523 <HER1>  
 A:Cross-references: EMBL:Z14118; NID:956863; PID:956864  
 A:Experimental source: strain Sprague Dawley  
 R;Herrén, B.; Weyer, K.A.; Rouge, M.; Loetscher, P.; Pech, M.  
 submitted to the EMBL Data Library, July 1992  
 A:Description: Cross-species conservation in sequence and function of PDGF ligands and receptors  
 A:Reference number: S25096  
 A:Accession: S25100  
 A:Molecule type: mRNA  
 A:Residues: 33-149, 'R', 151-518, 'T', 520-523 <HER2>  
 A:Cross-references: EMBL:Z14118; NID:956863; PID:956864  
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
 C:Keywords: ATP; autophosphorylation; glycoprotein; heterodimer; homodimer; phosphoprotein  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-1088/Product: platelet-derived growth factor receptor alpha #status predicted <MAT>  
 F:24-523/Domain: extracellular #status predicted <EXT>  
 F:41-101/Domain: immunoglobulin homology <IMM1>  
 F:142-190/Domain: immunoglobulin homology <IMM2>  
 F:227-291/Domain: immunoglobulin homology <IMM3>  
 F:427-502/Domain: immunoglobulin homology <IMM4>  
 F:524-547/Domain: transmembrane #status predicted <TM>  
 F:548-1088/Domain: intracellular #status predicted <INT>  
 F:590-956/Domain: protein kinase homology <KIN>  
 F:598-606/Region: protein kinase ATP-binding motif  
 F:48-99, 149-188, 234-289, 434-500/Disulfide bonds: #status predicted  
 F:75, 76, 88, 102, 178, 352, 358, 457, 467/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:626/Active site: Lys #status predicted  
 F:848/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 74.4%; Score 32; DB 1; Length 1088;  
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
 ||:||||  
 Db 449 CKDIKCC 455

RESULT 11  
 ACMSIT  
 N:Inositol 1,4,5-trisphosphate receptor - mouse  
 N:Alternate names: inositol-1,4,5-trisphosphate-binding protein P400; membrane-associated  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 02-Jul-1998  
 C:Accession: S04844; S06796; C34955  
 R:Furukuchi, T.; Yoshikawa, S.; Mikoshiba, K.  
 Nucleic Acids Res. 17, 5385-5386, 1989  
 A:Title: Nucleotide sequence of cDNA encoding P400 protein in the mouse cerebellum.  
 A:Reference number: S04844; MUID:893345101  
 A:Accession: S04844  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-2749 <FUR>  
 A:Cross-references: EMBL:X15373; NID:953568; PID:953569  
 A:Experimental source: strain ICR; cerebellum  
 R:Furukuchi, T.; Yoshikawa, S.; Miyawaki, A.; Wada, K.; Maeda, N.; Mikoshiba, K.  
 Nature 342, 32-38, 1989  
 A:Title: Primary structure and functional expression of the inositol 1,4,5-trisphosphate  
 A:Reference number: S06796; MUID:90044039  
 A:Accession: S06796  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-2749 <FUR>  
 A:Cross-references: EMBL:X15373; NID:953568; PID:953569

A:Experimental source: strain ICR; cerebellum  
 R;Nordquist, D.T.; Kozak, C.A.; Orr, H.T.  
 J. Neurosci. 8, 4780-4789, 1988  
 A:Title: cDNA cloning and characterization of three genes uniquely expressed in cerebellum  
 A:Reference number: A92975; MUID:89068131  
 A:Accession: C34955  
 A:Molecule type: mRNA  
 A:Residues: 2250-2674, 'P', 2676-2749 <NOR>  
 C:Superfamily: inositol-trisphosphate receptor  
 C:Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein  
 F:890-907/Domain: transmembrane #status predicted <TM1>  
 F:1960-1976/Domain: transmembrane #status predicted <TM2>  
 F:2276-2294/Domain: transmembrane #status predicted <TM3>  
 F:2308-2336/Domain: transmembrane #status predicted <TM4>  
 F:2334-2350/Domain: transmembrane #status predicted <TM5>  
 F:2352-2372/Domain: transmembrane #status predicted <TM6>  
 F:2391-2407/Domain: transmembrane #status predicted <TM7>  
 F:2440-2462/Domain: transmembrane #status predicted <TM8>  
 F:2570-2589/Domain: transmembrane #status predicted <TM9>  
 F:1588, 1755/Binding site: phosphate (Ser) (covalent) #status predicted  
 F:2475, 2503, 2622, 2710/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.4%; Score 32; DB 1; Length 2749;  
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKCC 7  
 ||:||||  
 Db 1321 KFIKCC 1326

RESULT 12  
 QOBE26  
 BZLF2 protein - human herpesvirus 4 (strain B95-8)  
 C:Species: human herpesvirus 4, Epstein-Barr virus  
 C:Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 05-Sep-1997  
 C:Accession: F43042; A03767; S33016  
 R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.  
 Mol. Biol. Med. 1, 21-45, 1983  
 A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr vi  
 A:Reference number: A93065; MUID:850357113  
 A:Accession: F43042  
 A:Molecule type: DNA  
 A:Residues: 1-223 <BAN>  
 A:Cross-references: EMBL:V01555; NID:959074; PID:91334876  
 R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.  
 Nature 310, 207-211, 1984  
 A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.  
 A:Reference number: A03794; MUID:84270667  
 A:Contents: annotation; protein coding region  
 C:Superfamily: human herpesvirus 4 BZLF2 protein

Query Match 74.4%; Score 32; DB 1; Length 223;  
 Best Local Similarity 71.4%; Pred. No. 37;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
 ||:||||  
 Db 208 CSFLKPC 214

RESULT 13  
 WMNV38  
 PE-38 protein - Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus  
 C:Species: Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus, OpMNVP  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 13-Sep-1997  
 C:Accession: A42191  
 R:Theilmann, D.A.; Stewart, S.  
 Virology 187, 97-106, 1992  
 A:Title: Tandemly repeated sequence at the 3' end of the IE-2 gene of the baculovirus  
 A:Reference number: A42191; MUID:92142537

Db 190 CEFLKRC 196

RESULT 7  
B71444  
hypothetical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
A:Variety: Columbia  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 12-Feb-1999  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks  
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel  
avanagh, T.; Hempel, S.; Kottler, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech  
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans  
C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal  
A:Reference number: A71400; MUID:98121113  
A:Accession: B71444  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Cross-references: GB:297343; NID:g2245073; PID:e327063; PID:g2245105  
C:Genetics:  
A:Map position: 4COP9-4G3845  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
F:133-189/Domain: homeobox homology <HOX>

Query Match 76.7%; Score 33; DB 2; Length 282;  
Best Local Similarity 57.1%; Pred. No. 29;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKRC 7  
II:II:II  
Db 196 CEYLARC 202

RESULT 8  
T00402  
homeodomain-zipper protein, ATHB-4 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 12-Feb-1999  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Mas  
submitted to the EMBL Data Library, July 1997  
A:Description: Arabidopsis thaliana chromosome II BAC T13E15 genomic sequence.  
A:Reference number: 214146  
A:Accession: T00402  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-318 <ROU>  
A:Cross-references: EMBL:AC002388; NID:g3420042; PID:g2344893  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Gene: ATHB-4  
A:Map position: II  
A:Introns: 68/1; 187/3; 214/2  
A:Note: T13E15.8  
C:Superfamily: homeobox homology  
F:161-217/Domain: homeobox homology <HOX>

Query Match 76.7%; Score 33; DB 2; Length 318;  
Best Local Similarity 57.1%; Pred. No. 33;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKRC 7  
II:II:II  
Db 224 CEYLKRC 230

RESULT 9  
PFHUGA  
platelet-derived growth factor receptor alpha precursor - human  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 18-Sep-1998  
C:Accession: A40162; A32941  
R:Matsui, T.; Heidaran, M.; Miki, T.; Popescu, N.; La Rochelle, W.; Kraus, M.; Pierce  
Science 243, 800-804, 1989  
A:Title: Isolation of a novel receptor cDNA establishes the existence of two PDGF rec  
A:Reference number: A40162; MUID:89130149  
A:Accession: A40162  
A:Molecule type: mRNA  
A:Residues: 1-1089 <MATS>  
A:Cross-references: GB:M21574; NID:g189733; PID:g189734  
R:Clackson-Welsh, L.; Eriksson, A.; Westmark, B.; Heldin, C.H.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4917-4921, 1989  
A:Title: cDNA cloning and expression of the human A-type platelet-derived growth fact  
A:Reference number: A32941; MUID:89296915  
A:Accession: A32941  
A:Molecule type: mRNA  
A:Residues: 1-1089 <CLA>  
A:Cross-references: GB:M22734; NID:g189725; PID:g189726  
C:Comment: The extracellular domain is predicted to include five immunoglobulin-like  
C:Genetics:  
A:Gene: GDB:PDGFRA  
A:Cross-references: GDB:I20267; OMIM:173490  
A:Map position: 4q11-q12  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP; autophosphorylation; dimer; glycoprotein; growth factor receptor; ph  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-1089/Product: platelet-derived growth factor receptor alpha #status predicted <M  
F:25-524/Domain: extracellular #status predicted <EXT>  
F:42-102/Domain: immunoglobulin homology <IMM1>  
F:143-191/Domain: immunoglobulin homology <IMM2>  
F:228-292/Domain: immunoglobulin homology <IMM3>  
F:428-503/Domain: immunoglobulin homology <IMM4>  
F:525-548/Domain: transmembrane #status predicted <TM>  
F:549-1089/Domain: intracellular #status predicted <INT>  
F:591-957/Domain: protein kinase homology <KIN>  
F:599-607/Region: protein kinase ATP-binding motif  
F:42,76,103,179,353,359,458,468/Binding site: carbohydrate (Asn) (covalent) #status p  
F:49-100,150-189,235-290,435-501/Disulfide bonds: #status predicted  
F:627/Active site: Lys #status predicted  
F:849/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

Query Match 74.4%; Score 32; DB 1; Length 1089;  
Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFLKRC 7  
II:II:II  
Db 450 CKDIKRC 456

RESULT 10  
PFRTGA  
platelet-derived growth factor receptor alpha precursor - rat  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 26-Feb-1999  
C:Accession: A34710; S33767; S25100  
R:Lee, K.H.; Bowen-Pope, D.F.; Reed, R.R.  
Mol. Cell. Biol. 10, 2237-2246, 1990  
A:Title: Isolation and characterization of the alpha platelet-derived growth factor r  
A:Reference number: A34710; MUID:90220609  
A:Accession: A34710  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1088 <LEE>  
A:Cross-references: GB:M63837; NID:g202929; PID:g202930

A:Title: Retinoic acid promotes transcription of the platelet-derived growth factor alpha  
A:Reference number: I57511; MUID:91061789  
A:Accession: I57511  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1089 <RES>  
A:Cross-references: GB:M57683; NID:g199783; PID:g199784  
C:Genetics:  
A:Gene: PDGF-alpha-R  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; growth factor receptor; transmembrane protein  
F:228-232/Domain: immunoglobulin homology <IMM1>  
F:591-957/Domain: protein kinase homology <KIN>  
F:599-607/Region: protein kinase ATP-binding motif

Query Match 79.1%; Score 34; DB 2; Length 1089;  
Best Local Similarity 71.4%; Pred. No. 60;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
|||  
Db 450 CKHIKC 456

RESULT 4  
G70485  
hypothetical protein aq\_2164 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 08-May-1998  
A:Accession: G70485  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98198666  
A:Accession: G70485  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-189 <AQF>  
A:Cross-references: GB:AE000776; NID:g2984355; PID:g2984368; GB:AE000657  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: aq\_2164

Query Match 79.1%; Score 34; DB 2; Length 189;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKCC 7  
|||||  
Db 131 KFLKCC 136

RESULT 5  
S31424  
DNA-binding homeotic protein Athb-2 - Arabidopsis thaliana  
N:Alternate names: DNA-binding protein Athb-2  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: columbia  
C:Accession: S31424; S28566; B71435  
C:Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 28-Aug-1998  
R:Carabelli, M.; Sessa, G.; Balma, S.; Morelli, G.; Ruberti, I.  
Plant J. 4, 469-479, 1993  
A:Title: The Arabidopsis Athb-2 and -4 genes are strongly induced by far-red-rich light.  
A:Reference number: S28565; MUID:94035194  
A:Note: Athb-2 and -4 are members of the HD-ZIPII protein family in Arabidopsis thaliana  
A:Accession: S31424  
A:Molecule type: DNA  
A:Residues: 1-284 <CAR>  
A:Cross-references: EMBL:X68146; NID:g16330; PID:g16332  
A:Accession: S28566

A:Molecule type: mRNA  
A:Residues: 1-284 <CAF>  
A:Cross-references: EMBL:X68145; NID:g16178; PID:g16180  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D  
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weissenegger, T.; Pohl, T.M.; Terry, N.; G  
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen  
erhoff, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;  
C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t  
A:Reference number: A71400; MUID:98121113  
A:Accession: B71435  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-284 <BEV>  
A:Cross-references: GB:297342; NID:g2245031; PID:e327013; PID:g2245033  
C:Genetics:  
A:Gene: Athb-2  
A:Map position: 4COP9-4G3845  
A:Introns: 54/1; 153/3; 180/2  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:127-183/Domain: homeobox homology <HOX>

Query Match 76.7%; Score 33; DB 2; Length 284;  
Best Local Similarity 57.1%; Pred. No. 30;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
|||||  
Db 190 CEFLRRC 196

RESULT 6  
A46304  
homeotic protein HAT4 - Arabidopsis thaliana  
N:Alternate names: developmental regulator HAT4  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 17-Oct-1997  
C:Accession: A46304; A44088; S27753; S29997  
R:Schen, M.; Lloyd, A.M.; Davis, R.W.  
Genes Dev. 7, 367-379, 1993  
A:Title: The HAT4 gene of Arabidopsis encodes a developmental regulator.  
A:Reference number: A46304; MUID:93194057  
A:Accession: A46304  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-284 <SCH>  
A:Cross-references: EMBL:Z19602; NID:g22758; PID:g22759  
A:Note: sequence extracted from NCBI backbone (NCBIN:127135, NCBIP:127137)  
R:Schen, M.; Davis, R.W.  
Proc. Natl. Acad. Sci. U.S.A. 89, 3894-3898, 1992  
A:Title: HD-zip proteins: members of an Arabidopsis homeodomain protein superfamily.  
A:Reference number: A44088; MUID:9223725  
A:Accession: A44088  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-284 <SCF>  
A:Cross-references: EMBL:M90394; NID:g166751; PID:g166752  
C:Genetics:  
A:Introns: 54/1; 153/3; 180/2  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:127-183/Domain: homeobox homology <HOX>

Query Match 76.7%; Score 33; DB 2; Length 284;  
Best Local Similarity 57.1%; Pred. No. 30;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:44 ; Search time 105.14 Seconds  
(without alignments)  
2.667 Million cell updates/sec

Title: US-09-124-280A-30

Perfect score: 43

Sequence: 1 CKFLKKC 7

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	90.7	73	2 C44088	homeotic protein H
2	37	86.0	239	2 T03078	conserved hypothetical
3	34	79.1	1089	2 S33727	platelet-derived g
4	34	79.1	189	2 G70485	hypothetical prote
5	33	76.7	284	2 S31424	DNA-binding homeot
6	33	76.7	284	2 A46304	homeotic protein H
7	33	76.7	282	2 B71444	hypothetical prote
8	33	76.7	318	2 T00402	homeodomain-zipper
9	32	74.4	1089	1 PFHUGA	platelet-derived g
10	32	74.4	1088	1 PRTGA	platelet-derived g
11	32	74.4	2749	1 ACMSIT	inositol 1,4,5-tri
12	32	74.4	223	1 QBE26	BLF2 protein - hu
13	32	74.4	307	1 WNV38	PE-38 protein - Or
14	32	74.4	1097	2 S47220	protein kinase C (
15	32	74.4	1151	2 S45390	protein kinase C (
16	32	74.4	24	2 JG1355	brevinin-1 - Frog
17	32	74.4	2734	2 B36579	inositol 1,4,5-tri
18	32	74.4	2749	2 A36579	inositol 1,4,5-tri
19	32	74.4	2713	2 A57113	inositol 1,4,5-tri
20	32	74.4	2695	2 S54974	type 1 inositol 1,
21	31	72.1	1115	1 IJMSNL	neural cell adhesi
22	31	72.1	858	1 IJRTNC	neural cell adhesi
23	31	72.1	853	1 IJBOHC	neural cell adhesi
24	31	72.1	1091	1 IJCHNL	neural cell adhesi
25	31	72.1	197	1 VHXRPD	minor outer capsid
26	31	72.1	355	2 B26883	neural cell adhesi
27	31	72.1	421	2 A60058	neural cell adhesi
28	31	72.1	244	2 A49202	cysteine-rich secr
29	31	72.1	816	2 S64439	hypothetical prote
30	31	72.1	2539	2 B71619	hypothetical prote
31	31	72.1	1700	2 S08167	Balbani ring 3 pr
32	31	72.1	311	2 S08110	salivary protein 1
33	31	72.1	984	2 S14382	hypothetical prote
34	31	72.1	403	2 S07825	hypothetical prote
35	30	69.8	538	1 O4CKA4	cytochrome P450 52
36	30	69.8	1147	1 MWXIB	myosin heavy chain
37	30	69.8	1092	1 JN0635	neural cell adhesi
38	30	69.8	1088	1 IJXLNL	neural cell adhesi
39	30	69.8	538	2 B40576	cytochrome P450 AL

ALIGNMENTS

RESULT 1

C44088

homeotic protein HAT22 - Arabidopsis thaliana (fragments)

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 02-Jul-1998

C:Accession: C44088

R:Scheda, M.; Davis, R.W.

Proc. Natl. Acad. Sci. U.S.A. 89, 3894-3898, 1992

A:Title: HD-Zip proteins: members of an Arabidopsis homeodomain protein superfamily.

A:Reference number: A44088; MUID:92237275

A:Accession: C44088

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-73 <SCH>

A:Cross-references: GB:M90417

C:Superfamily: unassigned homeobox proteins: homeobox homology

F:18-32/Domain: homeobox homology (fragment) <HOX>

Query Match 90.7%; Score 39; DB 2; Length 73;  
Best Local Similarity 85.7%; Pred. No. 0.73;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKKC 7

||:|||||

Db 39 CEFLKKC 45

RESULT 2

T03078

conserved hypothetical protein 061L - Chilo iridescent virus

C:Species: Chilo iridescent virus

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999

C:Accession: T03078

R:Bahr U.; Tidona, C.A.; Darai, G.

Virus Genes 15, 235-245, 1997

A:Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.

A:Reference number: Z14834

A:Accession: T03078

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-239 <BAH>

A:Cross-references: EMBL:AF003534; NID:g2738385; PID:g2738425

Query Match 86.0%; Score 37; DB 2; Length 239;

Best Local Similarity 71.4%; Pred. No. 4.6;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKKC 7

||:|||||

Db 87 CKYVKKC 93

RESULT 3

S33727

platelet-derived growth factor receptor alpha precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Apr-1998

C:Accession: I57511; S33727

R:Stiles, C.D.; Wang, C.

Mol. Cell. Biol. 10, 6781-6784, 1990

**THIS PAGE BLANK (USPTO)**

; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-31

Query Match 100.0%; Score 43; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKKC 7  
| | | | |  
Db 4 CKFLKKC 10

Search completed: September 7, 1999, 23:07:48  
Job time: 1746 sec



Qy 1 CKFLKCC 7  
| | | | |  
Db 4 CKFLKCC 10

## RESULT 13

US-08-653-632-50  
; Sequence 50, Application US/08653632  
; Patent No. 5830860  
; GENERAL INFORMATION:  
; APPLICANT: GRAY, Beulah  
; APPLICANT: HASEMAN, Judith R.  
; APPLICANT: MAYO, Kevin  
; TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
; STREET: 3100 No. 5830860west Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/653,632  
; FILING DATE: 24-MAY-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/218026  
; FILING DATE: 24-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kettelberger, Denise M  
; REGISTRATION NUMBER: 33,924  
; REFERENCE/DOCKET NUMBER: 600.286US11  
; TELEPHONE: 612/371-5268  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; US-08-653-632-50

Query Match 100.0%; Score 43; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.23; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
| | | | |  
Db 5 CKFLKCC 11

## RESULT 14

US-08-456-112B-30  
; Sequence 30, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porto, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-30

Query Match 100.0%; Score 43; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e-04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
| | | | |  
Db 1 CKFLKCC 7

## RESULT 15

US-08-456-112B-31  
; Sequence 31, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porto, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.

```

; STREET: 3100 No. 5786324west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,026
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 600.286US01
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-218-026-49

Query Match 100.0%; Score 43; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKKC 7
Db 4 CKFLKKC 10

RESULT 11
US-08-218-026-50
; Sequence 50, Application US/08218026
; Patent No. 5786324
; GENERAL INFORMATION:
; APPLICANT: Gray, Beulah
; APPLICANT: Haseman, Judith R.
; TITLE OF INVENTION: Synthetic Peptides with Bactericidal
; TITLE OF INVENTION: Activity and Endotoxin Neutralizing Activity for Gram
; TITLE OF INVENTION: Negative Bacteria and Methods for Their Use
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5786324west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,026
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 600.286US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300

Query Match 100.0%; Score 43; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKKC 7
Db 4 CKFLKKC 10

RESULT 11
US-08-218-026-50
; Sequence 50, Application US/08218026
; Patent No. 5786324
; GENERAL INFORMATION:
; APPLICANT: Gray, Beulah
; APPLICANT: Haseman, Judith R.
; TITLE OF INVENTION: Synthetic Peptides with Bactericidal
; TITLE OF INVENTION: Activity and Endotoxin Neutralizing Activity for Gram
; TITLE OF INVENTION: Negative Bacteria and Methods for Their Use
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5786324west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,026
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 600.286US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300

; STREET: 3100 No. 5786324west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,026
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 600.286US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300

; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
; US-08-218-026-50

Query Match 100.0%; Score 43; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKKC 7
Db 5 CKFLKKC 11

RESULT 12
US-08-653-632-49
; Sequence 49, Application US/08653632
; Patent No. 5830860
; GENERAL INFORMATION:
; APPLICANT: Gray, Beulah
; APPLICANT: HASEMAN, Judith R.
; APPLICANT: MAYO, Kevin
; TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 5830860west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/653,632
; FILING DATE: 24-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/218026
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 600.286US11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/371-5268
; TELEFAX: 612/332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-653-632-49

Query Match 100.0%; Score 43; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-1

Query Match 100.0%; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
Db 1 CKFLKCC 7

RESULT 8  
US-08-280-397-2  
Sequence 2, Application US/08280397  
Patent No. 5589459  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
of Bacterial Endotoxins and for the  
Prevention and Treatment of Septic  
Shock  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,397  
FILING DATE: 07/26/94  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-2

Query Match 100.0%; Score 43; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
Db 4 CKFLKCC 10

RESULT 9  
US-08-280-397-6  
Sequence 6, Application US/08280397  
Patent No. 5589459  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
of Bacterial Endotoxins and for the  
Prevention and Treatment of Septic  
Shock  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,397  
FILING DATE: 07/26/94  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-6

Query Match 100.0%; Score 43; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
Db 5 CKFLKCC 11

RESULT 10  
US-08-218-026-49  
Sequence 49, Application US/08218026  
Patent No. 5786324  
GENERAL INFORMATION:  
APPLICANT: Gray, Beulah  
APPLICANT: Haseman, Judith R.  
APPLICANT: Mayo, Kevin  
TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
Activity and Endotoxin Neutralizing Activity for Gram  
Negative Bacteria and Methods for Their Use  
TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould

Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
| | | | |  
Db 1 CKFLKCC 7

RESULT 5  
US-07-819-893-2  
; Sequence 2, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/819,893

FILING DATE: 19920115  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids

TYPE: AMINO ACID  
TOPOLOGY: circular

US-07-819-893-2

Query Match 100.0%; Score 43; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
| | | | |  
Db 4 CKFLKCC 10

RESULT 6  
US-07-819-893-6  
; Sequence 6, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/819,893  
FILING DATE: 19920115

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids

TYPE: AMINO ACID  
TOPOLOGY: circular

US-07-819-893-6

Query Match 100.0%; Score 43; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.23; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

Qy 1 CKFLKCC 7  
| | | | |  
Db 5 CKFLKCC 11

RESULT 7  
US-08-280-397-1  
; Sequence 1, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.

STREET: 1185 Avenue of the Americas  
CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/280,397  
FILING DATE: 07/26/94

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92

;  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-049-871-2

Query Match 100.0%; Score 43; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKKC 7  
Db 4 CKFLKKC 10

RESULT 3  
US-08-049-871-6  
; Sequence 6, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514

;  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-049-871-6

Query Match 100.0%; Score 43; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKKC 7  
Db 5 CKFLKKC 11

RESULT 4  
US-07-819-893-1  
; Sequence 1, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
; US-07-819-893-1

Query Match 100.0%; Score 43; DB 1; Length 7;

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:48 ; Search time 82.37 Seconds  
(without alignments)  
0.839 Million cell updates/sec

Title: US-09-124-280A-30  
Perfect score: 43  
Sequence: 1 CKFLKRC 7  
Scoring table: BLOSUM62  
Searched: 106577 seqs, 9868381 residues  
Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	7	1	US-08-049-871-1
2	43	100.0	10	1	US-08-049-871-2
3	43	100.0	11	1	US-08-049-871-6
4	43	100.0	7	1	US-07-819-893-1
5	43	100.0	10	1	US-07-819-893-2
6	43	100.0	11	1	US-07-819-893-6
7	43	100.0	7	1	US-08-280-397-1
8	43	100.0	10	1	US-08-280-397-2
9	43	100.0	11	1	US-08-280-397-6
10	43	100.0	10	2	US-08-218-026-49
11	43	100.0	11	2	US-08-218-026-50
12	43	100.0	10	2	US-08-653-632-49
13	43	100.0	11	2	US-08-653-632-50
14	43	100.0	7	2	US-08-456-112B-30
15	43	100.0	10	2	US-08-456-112B-31
16	43	100.0	11	2	US-08-456-112B-35
17	32	74.4	1089	1	US-08-180-195-36
18	32	74.4	1089	1	US-08-168-917-4
19	32	74.4	223	1	US-08-430-633-1
20	32	74.4	1089	2	US-08-477-329-36
21	32	74.4	1089	2	US-08-475-458-36
22	32	74.4	1089	2	US-08-460-510-4
23	32	74.4	1089	3	PCT-US92-00730-4
24	32	74.4	1089	3	PCT-US92-00862-4
25	32	74.4	2713	3	PCT-US96-01735-1
26	31	72.1	45	1	US-08-632-691-1
27	31	72.1	46	1	US-08-462-661A-13
28	31	72.1	46	1	US-08-462-661A-15
29	31	72.1	355	2	US-08-458-555-2
30	31	72.1	391	2	US-08-644-034A-1
31	30.5	70.9	12	1	US-08-097-830B-20
32	30.5	70.9	12	1	US-08-456-112B-20
33	30	69.8	49	1	US-07-662-225A-1
34	30	69.8	49	1	US-07-662-225A-2
35	30	69.8	49	1	US-07-662-225A-3
36	30	69.8	39	1	US-07-662-225A-4
37	30	69.8	39	1	US-07-662-225A-5
38	30	69.8	49	1	US-07-802-847C-22
39	30	69.8	98	1	US-07-623-611-6

40 30 69.8 98 1 US-07-623-611-9 Sequence 9, Appli  
41 30 69.8 49 1 US-07-965-674-6 Sequence 6, Appli  
42 30 69.8 98 3 PCT-US91-09108-6 Sequence 6, Appli  
43 30 69.8 98 3 PCT-US91-09108-9 Sequence 9, Appli  
44 30 69.8 49 3 PCT-US93-05640-21 Sequence 21, Appli  
45 30 69.8 49 3 PCT-US93-09523-6 Sequence 6, Appli

ALIGNMENTS

RESULT 1  
US-08-049-871-1  
; Sequence 1, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-1  
  
Query Match 100.0%; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CKFLKRC 7  
| | | | | | |  
Db 1 CKFLKRC 7  
  
RESULT 2  
US-08-049-871-2  
; Sequence 2, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the

DT 10-NOV-1998 (first entry)  
DE Streptococcus pneumoniae protein of unknown function.  
KW Streptococcus pneumoniae protein; genetic immunisation; antagonist;  
KW immunological response; inoculation; antibody production; inhibitor;  
KW T cell immune response; antimicrobial compound; bacterial adhesion;  
KW extracellular matrix protein; protein-mediated cell invasion; wound;  
KW pathogenesis.  
OS Streptococcus pneumoniae.  
PN W09743303-A1.  
PD 20-NOV-1997.  
PF 14-MAY-1997; U07950.  
PR 14-MAY-1996; US-017670.  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO,  
PI Stodola RK.  
DR WPI; 98-008793/01.  
DR N-PSDB; T98774.  
PT Novel Streptococcus pneumoniae proteins and related DNA - useful for  
PT diagnosing anti-microbial agents for treatment of bacterial  
PT infections  
PS Claim 12; Pages 460-461; 483pp; English.  
CC This sequence represents a Streptococcus pneumoniae protein of  
CC unknown function, and is encoded by a DNA sequence of the invention.  
CC The DNA sequences were isolated from Streptococcus pneumoniae strain  
CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the  
CC invention can be used to identify compounds which interact with and  
CC inhibit or activate the activity of the proteins. Antagonists can be  
CC used to treat diseases caused by S. pneumoniae proteins, through genetic  
CC immunisation. They can also be used to induce an immunological response  
CC in a mammal by inoculation with the S. pneumoniae proteins or delivery  
CC of the encoding nucleic acids in a vector adequate to produce antibody  
CC and/or T cell immune responses to protect the animal from disease. The  
CC proteins can also be used to identify antimicrobial compounds which are  
CC capable of inhibiting their bioactivity. In particular the proteins of  
CC the invention can be used to prevent adhesion of bacteria to mammalian  
CC extracellular matrix proteins on in-dwelling devices or in wounds, to  
CC block protein-mediated mammalian cell invasion, and to block the normal  
CC progression of pathogenesis in infections initiated other than by the  
CC implantation of in-dwelling devices or other surgical techniques.  
SQ Sequence 37 AA;

Query Match 79.1%; Score 34; DB 1; Length 37;  
Best Local Similarity 100.0%; Pred. No. 9.6;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KFLKKC 7  
Db 24 KFLKKC 29  
|||||

Search completed: September 7, 1999, 22:50:02  
Job time: 7883 sec

CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7 to  
 CC 10-9 M, and Escherichia coli. The peptides are also used to treat  
 CC endotoxin shock. The present sequence represents a peptide derived  
 CC from human neutrophil granule bactericidal protein from an example of  
 CC the present invention.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 43; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.11; 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;

QY 1 CKFLKCC 7  
 |||||  
 DB 5 CKFLKCC 11

RESULT 12  
 W01573  
 ID W01573 standard; Protein; 263 AA.  
 AC W01573;  
 DT 14-OCT-1997 (first entry)  
 DE Protein encoded by CH14-2a16-1, over abundant in cancer cells.  
 KW CH1-9a11-2; CH8-2a13-1; CH13-2a12-1; CH14-2a16-1; cancer; antibody;  
 KW diagnosis; treatment; breast cancer; ds.  
 OS Homo sapiens.  
 PN W09639516-A2.  
 PD 12-DEC-1996.  
 PF 05-JUN-1996; U09286.  
 PR 05-JUN-1996; US-463660.  
 PA (CALP-) CALIFORNIA PACIFIC MEDICAL RES INST.  
 PI Chen L, Smith H;  
 DR WPI: 97-043138/04.  
 DR N-PSDB: T58247.  
 PT Polynucleotide(s) over-abundant in cancer cells - used to develop  
 PT prods. for use in the diagnosis and treatment of cancer, partic.  
 PS breast cancer  
 PS Claim 18; Fig 20; 110pp; English.  
 CC The sequences given in W01569-75 are encoded by isolated polynucleotides,  
 CC selected from CH1-9a11-2, CH8-2a13-1, CH13-2a12-1, and CH14-2a16-1,  
 CC which are over-abundant in cancer cells. The DNA sequences, proteins  
 CC encoded by them, and antibodies directed against these proteins  
 CC can be used in the diagnosis and treatment of cancer, particularly  
 CC breast cancer. This sequence represents a protein which is suspected  
 CC of having DNA or RNA binding activity. This sequence is found to  
 CC be expressed at very high levels in testicular tissue.  
 SQ Sequence 263 AA;

Query Match 79.1%; Score 34; DB 1; Length 263;  
 Best Local Similarity 71.4%; Pred. No. 46;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFLKCC 7  
 |||||  
 DB 156 CKFAEKC 162

RESULT 13  
 W40374  
 ID W40374 standard; Protein; 263 AA.  
 AC W40374;  
 DT 21-JUL-1998 (first entry)  
 DE Human breast cancer protein CH14-2a16-1 from 2.0 kb DNA fragment.  
 KW Breast cancer; CH14-2a16-1; malignant transformation; diagnostic;  
 KW therapeutic; screening.  
 OS Homo sapiens.  
 PN W09738085-A2.  
 PD 16-OCT-1997.  
 PF 09-APR-1997; U05930.  
 PR 10-JUL-1996; US-678280.  
 PR 09-APR-1996; US-015167.  
 PR 05-JUN-1996; WO-009286.

PR 06-JUN-1996; US-019202.  
 PA (CALP-) CALIFORNIA PACIFIC MEDICAL CENT RES INST.  
 PI Chen L, Smith H;  
 DR WPI: 97-512705/47.  
 DR N-PSDB: V10693.  
 PT Breast cancer genes - used to develop products to design or screen  
 PT diagnostic reagents or therapeutic compounds  
 PS Claim 17; Fig 20; 118pp; English.  
 CC This sequence represents a novel human breast cancer protein CH14-2a16-1  
 CC isolated from a 2.0 kb DNA fragment. This gene fragment can be used for  
 CC identifying genes and gene products that are intimately related to  
 CC malignant transformation or maintenance of the malignant properties of  
 CC cancer cells. It can also be used to design or screen diagnostic reagents  
 CC or therapeutic compounds. Kits are included within the scope of the  
 CC invention. 263 AA;  
 SQ Sequence 263 AA;

Query Match 79.1%; Score 34; DB 1; Length 263;  
 Best Local Similarity 71.4%; Pred. No. 46;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFLKCC 7  
 |||||  
 DB 156 CKFAEKC 162

RESULT 14  
 W40378  
 ID W40378 standard; Protein; 579 AA.  
 AC W40378;  
 DT 21-JUL-1998 (first entry)  
 DE Human breast cancer protein CH14-2a16-1 from 2.0 kb DNA fragment #2.  
 KW Breast cancer; CH14-2a16-1; malignant transformation; diagnostic;  
 KW therapeutic; screening.  
 OS Homo sapiens.  
 PN W09738085-A2.  
 PD 16-OCT-1997.  
 PF 09-APR-1997; U05930.  
 PR 10-JUL-1996; US-678280.  
 PR 09-APR-1996; US-015167.  
 PR 05-JUN-1996; WO-009286.  
 PR 06-JUN-1996; US-019202.  
 PA (CALP-) CALIFORNIA PACIFIC MEDICAL CENT RES INST.  
 PI Chen L, Smith H;  
 DR WPI: 97-512705/47.  
 DR N-PSDB: V10697.  
 PT Breast cancer genes - used to develop products to design or screen  
 PT diagnostic reagents or therapeutic compounds  
 PS Claim 17; Fig 26; 118pp; English.  
 CC This sequence represents a novel human breast cancer protein CH14-2a16-1  
 CC isolated from a 2.0 kb DNA fragment. This gene fragment can be used for  
 CC identifying genes and gene products that are intimately related to  
 CC malignant transformation or maintenance of the malignant properties of  
 CC cancer cells. It can also be used to design or screen diagnostic reagents  
 CC or therapeutic compounds. Kits are included within the scope of the  
 CC invention. 579 AA;  
 SQ Sequence 579 AA;

Query Match 79.1%; Score 34; DB 1; Length 579;  
 Best Local Similarity 71.4%; Pred. No. 86;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFLKCC 7  
 |||||  
 DB 472 CKFAEKC 478

RESULT 15  
 W38732  
 ID W38732 standard; Protein; 37 AA.  
 AC W38732;



PT amino acid sequence - reduces dose of antibiotic required

PS Claim 35; Page 27; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to

CC potentiate the activity of an antibiotic when they are co-administered

CC with the antibiotic. Compositions containing these peptides are used

CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of

CC the outer bacterial membrane to the antibiotic, allowing a reduction

CC in the dose of antibiotic required by 10-90% of the normal dose for

CC in vivo or in vitro application. Any toxic side effects are

CC correspondingly reduced.

SO Sequence 10 AA;

Query Match 100.0%; Score 43; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.1;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
| | | | |

Db 4 CKFLKCC 10

RESULT 9

ID W21623 standard; peptide; 11 AA.

AC W21623;

DT 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #35.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.

OS Synthetic.

FH Key Location/Qualifiers

FT disulfide\_bond 5..11

PN WO9638163-A1.

PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;

DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic

PT amino acid sequence - reduces dose of antibiotic required

PS Claim 39; Page 28; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to

CC potentiate the activity of an antibiotic when they are co-administered

CC with the antibiotic. Compositions containing these peptides are used

CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of

CC the outer bacterial membrane to the antibiotic, allowing a reduction

CC in the dose of antibiotic required by 10-90% of the normal dose for

CC in vivo or in vitro application. Any toxic side effects are

CC correspondingly reduced.

SO Sequence 11 AA;

Query Match 100.0%; Score 43; DB 1; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
| | | | |

Db 5 CKFLKCC 11

RESULT 10

ID W62434 standard; peptide; 10 AA.

AC W62434;

DT 01-OCT-1998 (first entry)

DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:49.

KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;

KW bactericidal permeability increasing factor; B/PI; neutralising;

KW human neutrophil granule bactericidal protein.

OS Homo sapiens.

OS Synthetic.

PN US5786324-A.

PD 28-JUL-1998.

PF 24-MAR-1994; 218026.

PR 24-MAR-1994; US-218026.

PA (MINU ) UNIV MINNESOTA.

PI Gray B, Haseman JR, Mayo K;

DR WPI; 98-436578/37.

PT Bactericidal and endotoxin-neutralising peptides - used in treating

PT e.g. Pseudomonas species infection and in protectively coating

PT prosthetic devices

PS Example 1; Column 18; 46pp; English.

CC The present invention describes bactericidal peptides (BP): (a) for

CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)

CC with both endotoxin neutralising activity and bactericidal activity for

CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising

CC activity and bactericidal activity, the peptide having at least 1 amino

CC acid sequence of a peptide with endotoxin neutralising activity combined

CC with at least 1 different amino acid sequence of a peptide that has

CC bactericidal activity; and (e) a prosthetic device which has a

CC sufficient amount of BP attached to the surface to inhibit bacterial

CC growth. The peptides of the invention are used in treating bacterial

CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7 to

CC 10-9 M, and Escherichia coli. The peptides are also used to treat

CC endotoxin shock. The present sequence represents a peptide derived

CC from human neutrophil granule bactericidal protein from an example of

CC the present invention.

SO Sequence 10 AA;

Query Match 100.0%; Score 43; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.1;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFLKCC 7  
| | | | |

Db 4 CKFLKCC 10

RESULT 11

ID W62435 standard; peptide; 11 AA.

AC W62435;

DT 01-OCT-1998 (first entry)

DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:50.

KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;

KW bactericidal permeability increasing factor; B/PI; neutralising;

KW human neutrophil granule bactericidal protein.

OS Homo sapiens.

OS Synthetic.

PN US5786324-A.

PD 28-JUL-1998.

PF 24-MAR-1994; 218026.

PR 24-MAR-1994; US-218026.

PA (MINU ) UNIV MINNESOTA.

PI Gray B, Haseman JR, Mayo K;

DR WPI; 98-436578/37.

PT Bactericidal and endotoxin-neutralising peptides - used in treating

PT e.g. Pseudomonas species infection and in protectively coating

PT prosthetic devices

PS Example 1; Column 18; 46pp; English.

CC The present invention describes bactericidal peptides (BP): (a) for

CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)

CC with both endotoxin neutralising activity and bactericidal activity for

CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising

CC activity and bactericidal activity, the peptide having at least 1 amino

CC acid sequence of a peptide with endotoxin neutralising activity combined

CC with at least 1 different amino acid sequence of a peptide that has

CC bactericidal activity; and (e) a prosthetic device which has a

CC sufficient amount of BP attached to the surface to inhibit bacterial

CC growth. The peptides of the invention are used in treating bacterial

OS Synthetic.  
 FH Key  
 FT disulfide bond 4. .10  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M.  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 4; Page 32; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 43; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKKC 7  
 DB 4 CKFLKKC 10

RESULT 6  
 R39292  
 ID R39292 standard; peptide; 11 AA.  
 AC R39292; 11 AA.  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 FH Key  
 FT disulfide bond 5. .11  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M.  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 8; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.

SQ Sequence 11 AA;

Query Match 100.0%; Score 43; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKKC 7  
 DB 5 CKFLKKC 11

RESULT 7  
 W21618  
 ID W21618 standard; peptide; 7 AA.  
 AC W21618;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #30.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key  
 FT disulfide bond 1. .7  
 PN WO9638163-Al.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M. Varra M.  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 34; Page 27; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKKC 7  
 DB 1 CKFLKKC 7

RESULT 8  
 W21619  
 ID W21619 standard; peptide; 10 AA.  
 AC W21619;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #31.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key  
 FT disulfide bond 4. .10  
 PN WO9638163-Al.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M. Varra M.  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic

PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 4; Page 32; 39pp; English.  
CC This is a specific example of a generic peptide of formula  
CC R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)<sub>n</sub>-R, where n = 1-100  
CC and each R is H, an amino acid residue or a fatty acid residue.  
CC The peptide is useful for treating or preventing septic shock,  
CC mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 43; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CKFLKKC 7  
| | | | | | |  
DB 4 CKFLKKC 10

RESULT 3  
ID R33531 standard; peptide; 11 AA.  
AC R33531;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 5..11  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 8; Page 32; 39pp; English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)<sub>n</sub>-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 11 AA;

Query Match 100.0%; Score 43; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CKFLKKC 7  
| | | | | | |  
DB 5 CKFLKKC 11

RESULT 4  
ID R39287 standard; peptide; 7 AA.  
AC R39287;  
DT 22-DEC-1993 (first entry)  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
KW quantification.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..7  
PN WO9314115-A.  
PD 22-JUL-1993.  
PF 14-MAY-1992; E01060.  
PR 16-JAN-1992; US-819893.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI; 93-243143/30.  
PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
PT forming non-toxic, antigenic complex, used to treat or prevent  
PT septic shock, in vaccines to detoxify blood, etc.  
PS Claim 3; Page 32; 45pp; English.  
CC The sequence is that of a peptide which binds to the lipid A  
CC component of bacterial endotoxin at the same site as polymyxin B  
CC (PMB) and with about the same affinity to produce a non-toxic  
CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
CC proteolytic degradation in serum, has no antibiotic activity and no  
CC haemolytic action. It is especially used to treat or prevent septic  
CC shock, to reduce toxicity of PMB; to remove endotoxins from blood,  
CC sera, vaccines, drug solns. etc.; to control release of cytokines,  
CC induced by endotoxins; for in vivo or in vitro detoxification of  
CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
CC products.  
SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CKFLKKC 7  
| | | | | | |  
DB 1 CKFLKKC 7

RESULT 5  
ID R39288 standard; peptide; 10 AA.  
AC R39288;  
DT 22-DEC-1993 (first entry)  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
KW quantification.

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:01 ; Search time 135.78 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-30

Perfect score: 43

Sequence: 1 CKFLKKC 7

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	7	1 R3325	Peptide for treati
2	43	100.0	10	1 R3326	Peptide for treati
3	43	100.0	11	1 R3331	Peptide for treati
4	43	100.0	7	1 R39287	Endotoxin lipid A
5	43	100.0	10	1 R39288	Endotoxin lipid A
6	43	100.0	11	1 R39292	Endotoxin lipid A
7	43	100.0	7	1 W21618	Antibiotic potenti
8	43	100.0	10	1 W21619	Antibiotic potenti
9	43	100.0	11	1 W21623	Antibiotic potenti
10	43	100.0	11	1 W62434	Human neutrophil g
11	43	100.0	10	1 W62435	Human neutrophil g
12	34	79.1	263	1 W01573	Protein encoded by
13	34	79.1	263	1 W40374	Human breast cance
14	34	79.1	579	1 W40378	Human breast cance
15	34	79.1	37	1 W38732	Streptococcus pneu
16	34	79.1	18	1 W99423	Cationic peptide B
17	34	79.1	17	1 W99424	Cationic peptide B
18	34	79.1	173	1 Y07084	Renal cancer assoc
19	32	74.4	1089	1 R06910	Alpha type PDGF re
20	32	74.4	1089	1 R08267	Platelet derived g
21	32	74.4	2749	1 R13887	Inositol-3-phospha
22	32	74.4	1009	1 R26206	Type B human plate
23	32	74.4	24	1 R51039	Antibacterial pept
24	32	74.4	223	1 R87020	BZLF2 protein, Eps
25	32	74.4	2713	1 W00168	Human T cell inosi
26	32	74.4	2695	1 W07632	Human type I inosi
27	32	74.4	223	1 W47351	Epstein-Barr virus
28	31	72.1	46	1 R10225	Platelet aggregati
29	31	72.1	58	1 R42857	Platelet aggregati
30	31	72.1	58	1 R42859	Platelet aggregati
31	31	72.1	58	1 R42860	Platelet aggregati
32	31	72.1	58	1 R42861	Platelet aggregati
33	31	72.1	58	1 R42862	Platelet aggregati
34	31	72.1	58	1 R46946	PAGalpha1-101 plat
35	31	72.1	58	1 R46948	PAGalpha3-101 plat
36	31	72.1	58	1 R46949	PAGalpha4-101 plat
37	31	72.1	58	1 R46950	PAGalpha5-101 plat
38	31	72.1	58	1 R46951	PAGalpha6-101 plat
39	31	72.1	383	1 R62083	Lactococcus lactis
40	31	72.1	72	1 W50453	Snake venom lactel
41	31	72.1	74	1 W50458	Snake venom platel
42	31	72.1	95	1 W50459	Snake venom PAI Da
43	31	72.1	67	1 W50460	Truncated snake ve

ALIGNMENTS

RESULT 1

R33525 R33525 standard; peptide; 7 AA.  
AC R33525;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..7  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992: 000943.  
PR 11-FEB-1991: US-658744.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 3; Page 32; 39pp; English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC µg/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFLKKC 7  
Db 1 CKFLKKC 7

RESULT 2

R33526 R33526 standard; peptide; 10 AA.  
AC R33526;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 4..10  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992: 000943.



QC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MARAS/VE90;  
RX MEDLINE; 9728302.  
RA RICO-HESE R., HARRISON L.M., SALAS R.A., TOVAR D., NISALAK A.,  
RA RAMOS C., BOSHELL J., DE MESA M.T., NOGUEIRA R.M., DA ROSA A.T.;  
RT "Origins of dengue type 2 viruses associated with increased  
RT pathogenicity in the Americas."  
RL Virology 230:244-251(1997).  
DR EMBL; U91866; AAC02036.1; -.  
KW Polyprotein.  
FT NON\_TER 1  
FT NON\_TER 80  
SQ SEQUENCE 80 AA; 8651 MW; 17ED2A0B CRC32;

Query Match 59.2%; Score 45; DB 12; Length 80;  
Best Local Similarity 60.0%; Pred. No. 0.4;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Caps 0;

Oy 3 SWKSSEIRCG 12  
|||: |||:  
Db 44 SWKNKELKCG 53

Search completed: September 7, 1999, 22:47:39  
Job time: 7959 sec

```

RESULT 11
O57086 PRELIMINARY; PRT; 80 AA.
AC O57086; 1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KD92-097/THA192;
RA RICO-HESE R., HARRISON L.M., NISALAK A., VAUGHN D.W., GREEN S.,
RA ROTHMAN A.L., ENNIS F.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87338; AAC02134.1; -.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 8751 MW; 2BD651C2 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 80;
Best Local Similarity 60.0%; Pred. No. 0.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
DB 44 SWKNKELKCG 53

RESULT 12
O57087 PRELIMINARY; PRT; 80 AA.
AC O57087; 1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D81-081/THA181;
RA RICO-HESE R., HARRISON L.M., NISALAK A., VAUGHN D.W., GREEN S.,
RA ROTHMAN A.L., ENNIS F.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87340; AAC02136.1; -.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 8751 MW; 2BD651C2 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 80;
Best Local Similarity 60.0%; Pred. No. 0.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
DB 44 SWKNKELKCG 53

RESULT 13
O57088 PRELIMINARY; PRT; 80 AA.
AC O57088; 1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)

```

```

DE POLYPROTEIN (FRAGMENT).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D82-137/THA182;
RA RICO-HESE R., HARRISON L.M., NISALAK A., VAUGHN D.W., GREEN S.,
RA ROTHMAN A.L., ENNIS F.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87341; AAC02137.1; -.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 8761 MW; 7199A0A9 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 80;
Best Local Similarity 60.0%; Pred. No. 0.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
DB 44 SWKNKELKCG 53

RESULT 14
O86855 PRELIMINARY; PRT; 127 AA.
AC O86855;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
GN E/NS1.
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95208980.
RA GUZMAN M.G., DEUBEL V., PELEGRINO J.L., ROSARIO D., MARRERO M.,
RA SARIOL C., KOURI G.;
RT "Partial nucleotide and amino acid sequences of the envelope and the
RT envelope/nonstructural protein-1 gene junction of four dengue-2 virus
RT strains isolated during the 1981 Cuban epidemic.";
RL An. J. Trop. Med. Hyg. 52:241-246(1995).
DR EMBL; S75632; AAB33854.1; -.
DR EMBL; S75630; AAB33854.1; JOINED.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 127 AA; 14004 MW; D9DE9CC6 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 127;
Best Local Similarity 60.0%; Pred. No. 0.62;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
DB 104 SWKNKELKCG 113

RESULT 15
O57127 PRELIMINARY; PRT; 80 AA.
AC O57127;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

```

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Query Match 59.2%; Score 45; DB 12; Length 150;  
Best Local Similarity 60.0%; Pred. No. 0.72;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
| | | : | : | |  
Db 58 SWKNKELKCG 67

RESULT 7  
O36691 PRELIMINARY; PRT; 150 AA.  
AC O36691;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
DE POLYPROTEIN (FRAGMENTS).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TSM;  
RA POIDINGER M., MACKENZIE J.S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF013773; AAB/0475.1; -.  
KW POLYPROTEIN.  
FT NON\_TER 1  
FT CHAIN 1  
FT CHAIN 1  
FT NON\_CONS 51  
FT CHAIN 52  
FT CHAIN 52  
FT NON\_TER 150  
FT NON\_TER 150  
SQ SEQUENCE 150 AA; 16413 MW; 22956552 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 150;  
Best Local Similarity 60.0%; Pred. No. 0.72;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
| | | : | : | |  
Db 58 SWKNKELKCG 67

RESULT 8  
O36692 PRELIMINARY; PRT; 150 AA.  
AC O36692;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
DE POLYPROTEIN (FRAGMENTS).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TSV01;  
RA POIDINGER M., MACKENZIE J.S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF013774; AAB/0476.1; -.  
KW POLYPROTEIN.  
FT NON\_TER 1  
FT CHAIN 1  
FT NON\_CONS 51  
FT CHAIN 52  
FT CHAIN 52  
FT NON\_TER 150  
FT NON\_TER 150  
SQ SEQUENCE 150 AA; 16453 MW; 3018ADF9 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 150;  
Best Local Similarity 60.0%; Pred. No. 0.72;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
| | | : | : | |  
Db 58 SWKNKELKCG 67

RESULT 9  
O57080 PRELIMINARY; PRT; 80 AA.  
AC O57080;  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
DE POLYPROTEIN (FRAGMENT).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-D86-337/THA186;  
RA RICO-HESE R., HARRISON L.M., NISALAK A., VAUGHN D.W., GREEN S.,  
RA ROHMAN A.L., ENNIS F.A.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U87325; AAC02121.1; -.  
KW POLYPROTEIN.  
FT NON\_TER 1  
FT NON\_TER 80  
FT NON\_TER 80  
SQ SEQUENCE 80 AA; 8677 MW; 80508B98 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 80;  
Best Local Similarity 60.0%; Pred. No. 0.4;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
| | | : | : | |  
Db 44 SWKNKELKCG 53

RESULT 10  
O57081 PRELIMINARY; PRT; 80 AA.  
AC O57081;  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
DE POLYPROTEIN (FRAGMENT).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-D87-1372/THA187;  
RA RICO-HESE R., HARRISON L.M., NISALAK A., VAUGHN D.W., GREEN S.,  
RA ROHMAN A.L., ENNIS F.A.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U87326; AAC02122.1; -.  
KW POLYPROTEIN.  
FT NON\_TER 1  
FT NON\_TER 80  
FT NON\_TER 80  
SQ SEQUENCE 80 AA; 8719 MW; 333D4E69 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 80;  
Best Local Similarity 60.0%; Pred. No. 0.4;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
| | | : | : | |  
Db 44 SWKNKELKCG 53



DR EMBL; U49189; AAA91809.1; -.  
KW Polyprotein.  
FT NON\_TER 1  
FT NON\_TER 80  
SQ SEQUENCE 80 AA; 8705 MW; 13F52871 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 80;  
Best Local Similarity 60.0%; Pred. No. 0.4;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SWKSSEIRCG 12  
|||: |::||  
Db 44 SWKNKELKCG 53

RESULT 3  
O36687 PRELIMINARY; PRT; 150 AA.

AC O36687; (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DE POLYPROTEIN (FRAGMENTS).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TSK;  
RA POIDINGER M., MACKENZIE J.S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF013769; AAB70471.1; -.  
KW Polyprotein.  
FT NON\_TER 1  
FT CHAIN 1 >51 ENVELOPE PROTEIN.  
FT NON\_CONS 51 52  
FT CHAIN 52 >150 NONSTRUCTURAL PROTEIN 1.  
FT NON\_TER 150 150  
SQ SEQUENCE 150 AA; 16413 MW; 22956552 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 150;  
Best Local Similarity 60.0%; Pred. No. 0.72;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SWKSSEIRCG 12  
|||: |::||  
Db 58 SWKNKELKCG 67

RESULT 4  
O36688 PRELIMINARY; PRT; 150 AA.

AC O36688; (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DE POLYPROTEIN (FRAGMENTS).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TSN;  
RA POIDINGER M., MACKENZIE J.S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF013770; AAB70472.1; -.  
KW Polyprotein.  
FT NON\_TER 1  
FT CHAIN 1 >51 ENVELOPE PROTEIN.  
FT NON\_CONS 51 52  
FT CHAIN 52 >150 NONSTRUCTURAL PROTEIN 1.

FT NON\_TER 150  
SQ SEQUENCE 150 AA; 16413 MW; 22956552 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 150;  
Best Local Similarity 60.0%; Pred. No. 0.72;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SWKSSEIRCG 12  
|||: |::||  
Db 58 SWKNKELKCG 67

RESULT 5  
O36689 PRELIMINARY; PRT; 150 AA.

AC O36689; (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DE POLYPROTEIN (FRAGMENTS).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TSM;  
RA POIDINGER M., MACKENZIE J.S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF013771; AAB70473.1; -.  
KW Polyprotein.  
FT NON\_TER 1  
FT CHAIN 1 >51 ENVELOPE PROTEIN.  
FT NON\_CONS 51 52  
FT CHAIN 52 >150 NONSTRUCTURAL PROTEIN 1.  
FT NON\_TER 150 150  
SQ SEQUENCE 150 AA; 16413 MW; 22956552 CRC32;

Query Match 59.2%; Score 45; DB 12; Length 150;  
Best Local Similarity 60.0%; Pred. No. 0.72;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SWKSSEIRCG 12  
|||: |::||  
Db 58 SWKNKELKCG 67

RESULT 6  
O36690 PRELIMINARY; PRT; 150 AA.

AC O36690; (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DE POLYPROTEIN (FRAGMENTS).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TSS;  
RA POIDINGER M., MACKENZIE J.S.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF013772; AAB70474.1; -.  
KW Polyprotein.  
FT NON\_TER 1  
FT CHAIN 1 >51 ENVELOPE PROTEIN.  
FT NON\_CONS 51 52  
FT CHAIN 52 >150 NONSTRUCTURAL PROTEIN 1.  
FT NON\_TER 150 150  
SQ SEQUENCE 150 AA; 16413 MW; 22956552 CRC32;

GenCore version 4.5

Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:38 ; Search time 160.2 Seconds  
(without alignments)  
4.994 Million cell updates/sec

Title: US-09-124-280A-29

Perfect score: 76

Sequence: 1 CQSWKSSEIRCGK 13

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp rodent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	136	5 Q25387	Q25387 limulus pol
2	45	59.2	80	12 Q66465	Q66465 dengue viru
3	45	59.2	150	12 Q36687	Q36687 dengue viru
4	45	59.2	150	12 Q36688	Q36688 dengue viru
5	45	59.2	150	12 Q36689	Q36689 dengue viru
6	45	59.2	150	12 Q36690	Q36690 dengue viru
7	45	59.2	150	12 Q36691	Q36691 dengue viru
8	45	59.2	150	12 Q36692	Q36692 dengue viru
9	45	59.2	80	12 Q57080	Q57080 dengue viru
10	45	59.2	80	12 Q57081	Q57081 dengue viru
11	45	59.2	80	12 Q57086	Q57086 dengue viru
12	45	59.2	80	12 Q57087	Q57087 dengue viru
13	45	59.2	80	12 Q57088	Q57088 dengue viru
14	45	59.2	127	12 Q86855	Q86855 dengue viru
15	45	59.2	80	12 Q57127	Q57127 dengue viru
16	45	59.2	80	12 Q57128	Q57128 dengue viru
17	45	59.2	80	12 Q57129	Q57129 dengue viru
18	45	59.2	80	12 Q57130	Q57130 dengue viru
19	45	59.2	80	12 Q57131	Q57131 dengue viru
20	45	59.2	80	12 Q57132	Q57132 dengue viru
21	45	59.2	80	12 Q57133	Q57133 dengue viru
22	45	59.2	352	12 Q04035	Q04035 dengue viru
23	45	59.2	352	12 Q04036	Q04036 dengue viru
24	45	59.2	352	12 Q04037	Q04037 dengue viru
25	45	59.2	352	12 Q04038	Q04038 dengue viru
26	45	59.2	352	12 Q04039	Q04039 dengue viru
27	45	59.2	352	12 Q04040	Q04040 dengue viru
28	45	59.2	352	12 Q04041	Q04041 dengue viru
29	45	59.2	352	12 Q04042	Q04042 dengue viru

30	45	59.2	352	12 Q66345	Q66345 dengue viru
31	45	59.2	1127	12 P89531	P89531 dengue viru
32	45	59.2	1127	12 P89532	P89532 dengue viru
33	45	59.2	578	12 O12290	O12290 dengue viru
34	45	59.2	577	12 O12291	O12291 dengue viru
35	45	59.2	352	12 Q67429	Q67429 dengue viru
36	45	59.2	352	12 Q67430	Q67430 dengue viru
37	45	59.2	1127	12 Q66454	Q66454 dengue viru
38	45	59.2	1127	12 Q66455	Q66455 dengue viru
39	45	59.2	1127	12 Q66456	Q66456 dengue viru
40	45	59.2	1127	12 Q66457	Q66457 dengue viru
41	45	59.2	3391	12 O11875	O11875 dengue viru
42	45	59.2	3391	12 O09234	O09234 unidentified
43	45	59.2	380	12 Q06371	Q06371 dengue viru
44	45	59.2	3391	12 Q92752	Q92752 dengue viru
45	45	59.2	3391	12 Q9YKL3	Q9YKL3 dengue viru

## ALIGNMENTS

```

RESULT 1
Q25387 ID Q25387 PRELIMINARY; PRT; 136 AA.
AC Q25387;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DE 01-NOV-1998 (TREMELrel. 08, Last annotation update)
DE ENDOTOXIN-BINDING PROTEIN-PROTEASE INHIBITOR PRECURSOR.
GN LEBP-PI.
OS Limulus polyphemus (Atlantic horseshoe crab).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC Limulidae; Limulus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92041932.
RA MINETTI C.A., LIN Y., CISLO T., LIU T.Y.;
RT "Purification and characterization of an endotoxin-binding protein
RT with protease inhibitory activity from Limulus ameobocytes.";
RL J. Biol. Chem. 268:20773-20780(1991).
DR EMBL; M65017; AAA28271.1; -.
KW Signal; Protease.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 136 ENDOTOXIN-BINDING PROTEIN-PROTEASE
FT INHIBITOR.
SQ SEQUENCE 136 AA; 15316 MW; 2ED8A280 CRC32;

Query Match 100.0%; Score 76; DB 5; Length 136;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQSWKSSEIRCGK 13
| | | | | | | | | |
DB 36 CQSWKSSEIRCGK 48

RESULT 2
Q66465 ID Q66465 PRELIMINARY; PRT; 80 AA.
AC Q66465;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DE 01-AUG-1998 (TREMELrel. 07, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PKKH-77/94;
RYAW-ZIN-THANT IGARASHI A.;
RA Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

```



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; M23027; G511851; -.  
DR PIR; A27032; GNMWVP.  
DR PFAM; PF00869; Flavi\_glycoprot; 1.  
DR PFAM; PF00948; Flavi\_NSI; 1.  
DR PFAM; PF01003; Flavi\_Capsid; 1.  
DR PFAM; PF01004; Flavi\_M; 1.  
DR PFAM; PF01005; Flavi\_NS2A; 1.  
DR HSP; P14336; ISVB.  
KW POLYPROTEIN; GLYCOPROTEIN; CORE PROTEIN; COAT PROTEIN;  
KW ENVELOPE PROTEIN; TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.  
FT CHAIN 1 114  
CAPSID PROTEIN C.  
FT PROPEP 115 205  
ENVELOPE GLYCOPROTEIN M.  
FT CHAIN 206 280  
MAJOR ENVELOPE PROTEIN E.  
FT CHAIN 281 775  
NONSTRUCTURAL PROTEIN NS1.  
FT CHAIN 776 1127  
NONSTRUCTURAL PROTEIN NS2A.  
FT CHAIN 1128 >1226  
POTENTIAL.  
FT TRANSMEM 46 67  
POTENTIAL.  
FT TRANSMEM 102 118  
POTENTIAL.  
FT TRANSMEM 244 263  
POTENTIAL.  
FT TRANSMEM 266 281  
POTENTIAL.  
FT TRANSMEM 715 735  
POTENTIAL.  
FT TRANSMEM 755 773  
POTENTIAL.  
FT DISULFID 283 310  
BY SIMILARITY.  
FT DISULFID 340 396  
BY SIMILARITY.  
FT DISULFID 354 385  
BY SIMILARITY.  
FT DISULFID 372 401  
BY SIMILARITY.  
FT DISULFID 465 565  
BY SIMILARITY.  
FT DISULFID 582 613  
BY SIMILARITY.  
FT CARBOHYD 183 183  
POTENTIAL.  
FT CARBOHYD 347 347  
POTENTIAL.  
FT CARBOHYD 433 433  
POTENTIAL.  
FT CARBOHYD 905 905  
POTENTIAL.  
FT CARBOHYD 982 982  
POTENTIAL.  
FT CARBOHYD 1190 1190  
POTENTIAL.  
FT NON\_TER 1226 1226  
SQ SEQUENCE 1226 AA; 136237 MW; 754BF3D4 CRC32;

Query Match 52.6%; Score 40; DB 1; Length 1226;  
Best Local Similarity 50.0%; Pred. No. 35;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
Db 782 NWKRELKCG 791

Search completed: September 7, 1999, 23:59:10  
Job time: 522 sec

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M73998; G164855; -  
 DR PIR; JQ1171; JQ1171;  
 DR PDB; 1LYP; 31-MAR-95.  
 DR PROSITE; PS00946; CATHELICIDINS\_1; 1.  
 DR PROSITE; PS00947; CATHELICIDINS\_2; 1.  
 DR PFAM; PF00666; Cathelicidins; 1.  
 KW ANTI-BIOTIC; SIGNAL; 3D-STRUCTURE.  
 FT SIGNAL 1 29 POTENTIAL.  
 FT CHAIN 30 171 ANTIMICROBIAL PROTEIN CAP18.  
 FT CHAIN 135 171 ANTIMICROBIAL PROTEIN CAP7.  
 FT MOD\_RES 30 30 PYRROLIDONE CARBOXYLIC ACID  
 (BY SIMILARITY).  
 FT DISULFID 85 96 BY SIMILARITY.  
 FT DISULFID 107 124 BY SIMILARITY.  
 FT VARIANT 157 157 K -> D.  
 SQ SEQUENCE 171 AA; 19805 MW; AC9639DA CRC32;

Query Match 52.6%; Score 40; DB 1; Length 171;  
 Best Local Similarity 50.0%; Pred. No. 5.1;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 QWSKSEIRCGK 13

Db 115 QAWDSFDICNR 126

RESULT 14

ID POLG\_DEN1S STANDARD; PRT; 3396 AA.  
 AC P33478;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX  
 DE PROTEIN (ENVELOPE PROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL  
 DE PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED  
 DE RNA POLYMERASE (EC 2.7.7.48) (NS5)].  
 OS DENGUE VIRUS TYPE 1 (STRAIN SINGAPORE S275/90).  
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;  
 OC FLAVIVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92263809.  
 RA FU J., TAN B.H., YAP E.H., CHAN Y.C., TAN Y.H.;  
 RT "Full-length cDNA sequence of dengue type 1 virus (Singapore strain  
 RT S275/90)."  
 RL VIROLOGY 188:953-958(1992).  
 CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF THREE PROTEINS:  
 CC PROTEIN PM, PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A  
 CC COMPLEX OF PROTEIN C AND MRNA.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M87512; -; NOT\_ANNOTATED\_CDS.

DR PIR; A42551; A42551.

DR PFAM; PF00869; Flavi\_glycoprot; 1.

DR PFAM; PF00948; Flavi\_NSI; 1.

DR PFAM; PF00949; Flavi\_helicase; 1.

DR PFAM; PF00972; Flavi\_NS5; 1.  
 DR PFAM; PF01002; Flavi\_NS2B; 1.  
 DR PFAM; PF01003; Flavi\_capsid; 1.  
 DR PFAM; PF01004; Flavi\_M; 1.  
 DR PFAM; PF01005; Flavi\_NS2A; 1.  
 DR HSP; P14336; ISVB.  
 KW POLYPROTEIN; GLYCOPROTEIN; TRANSFERASE; RNA-DIRECTED RNA POLYMERASE;  
 KW CORE PROTEIN; COAT PROTEIN; ENVELOPE PROTEIN; HELICASE; ATP-BINDING;  
 KW TRANSMEMBRANE; NONSTRUCTURAL CAPSID PROTEIN C (POTENTIAL).  
 FT CHAIN 1 114  
 FT PROPEP 115 205 ENVELOPE GLYCOPROTEIN M (POTENTIAL).  
 FT CHAIN 206 280 MAJOR ENVELOPE PROTEIN E (POTENTIAL).  
 FT CHAIN 281 774  
 FT CHAIN 775 1127 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).  
 FT CHAIN 1128 1344 NONSTRUCTURAL PROTEIN NS2A (POTENTIAL).  
 FT CHAIN 1345 1474 NONSTRUCTURAL PROTEIN NS2B (POTENTIAL).  
 FT CHAIN 1475 2093 HELICASE (NS3) (POTENTIAL).  
 FT CHAIN 2094 2243 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).  
 FT CHAIN 2244 2492 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).  
 FT CHAIN 2493 3396 RNA-DIRECTED RNA POLYMERASE (NS5)  
 (POTENTIAL).  
 FT NP\_BIND 1668 1675 ATP (POTENTIAL).  
 FT SITE 1759 1762 DEAH BOX.  
 FT TRANSMEM 267 279 POTENTIAL.  
 FT TRANSMEM 753 769  
 FT DISULFID 283 310 BY SIMILARITY.  
 FT DISULFID 340 396 BY SIMILARITY.  
 FT DISULFID 354 385 BY SIMILARITY.  
 FT DISULFID 372 401 BY SIMILARITY.  
 FT DISULFID 465 565 BY SIMILARITY.  
 FT DISULFID 582 613 BY SIMILARITY.  
 FT CARBOHYD 183 183 POTENTIAL.  
 FT CARBOHYD 347 347 POTENTIAL.  
 FT CARBOHYD 433 433 POTENTIAL.  
 SQ SEQUENCE 3396 AA; 379558 MW; F538880D CRC32;

Query Match 52.6%; Score 40; DB 1; Length 3396;  
 Best Local Similarity 50.0%; Pred. No. 95;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 SWKSSEIRCG 12

Db 781 NWKGRELKCG 790

RESULT 15

ID POLG\_DEN1W STANDARD; PRT; 1226 AA.  
 AC P17763;  
 DT 01-AUG-1990 (REL. 15, CREATED)  
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX  
 DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;  
 DE NONSTRUCTURAL PROTEINS NS1 AND NS2A] (FRAGMENT).  
 OS DENGUE VIRUS TYPE 1 (STRAIN WESTERN PACIFIC).  
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;  
 OC FLAVIVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88044504.  
 RA MASON P.W., MCADA P.C., MASON T.L., FOURNIER M.J.;  
 RT "Sequence of the dengue-1 virus genome in the region encoding the  
 RT three structural proteins and the major nonstructural protein NS1.";  
 RL VIROLOGY 161:262-267(1987).  
 CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:  
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  
 CC PROTEIN C AND MRNA.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

FT CARBOHYD 1661 1661 POTENTIAL.
FT CARBOHYD 2300 2300 POTENTIAL.
FT CARBOHYD 2304 2304 POTENTIAL.
FT CARBOHYD 2456 2456 POTENTIAL.
FT CARBOHYD 2702 2702 POTENTIAL.
FT CARBOHYD 2712 2712 POTENTIAL.
SQ SEQUENCE 3390 AA; 378057 MW; 6E76AD72 CRC32;

Query Match 53.9%; Score 41; DB 1; Length 3390;
Best Local Similarity 50.0%; Pred. NO. 65;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
DB 780 NWKGKELKCG 789

RESULT 12
2SS3 ARATH STANDARD; PRT; 164 AA.
ID 2SS3 ARATH STANDARD; PRT; 164 AA.
AC P15459;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE 2S SEED STORAGE PROTEIN 3 PRECURSOR (2S ALBUMIN STORAGE PROTEIN).
GN AT2S3
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; TRACHEOPHYTA;
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-CV. C24;
RC KREBBERS E.; HERDIES L.; DE CLERCQ A.; SEURINCK J.; LEEMANS J.;
RA VAN DAMME J.; SEGURA M.; GHEYSEN G.; VAN MONTAGU M.;
RA VANDERCKHOVE J.;
RT "Determination of the processing sites of an Arabidopsis 2S albumin
and characterization of the complete gene family.";
RL PLANT PHYSIOL. 87:859-866(1988).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-CV. C24;
RC CONCEICAO A.D.S.; KREBBERS E.;
RA SUBMITTED (JUL-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
RN SEQUENCE OF 103-164 FROM N.A.
RP STRAIN-CV. COLUMBIA;
RA RAYNAL M.; GRELLER F.; LAUDIE M.; MEYER Y.; COOKE R.; DELSENY M.;
RL SUBMITTED (OCT-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC -!- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
CC LINKED BY DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22035; G166616; -
DR EMBL; 224744; G395201; -
DR EMBL; 217580; G18724; -
DR PIR; JA0163; NWM03.
DR PIR; S34674; S34674.
KW SEED STORAGE PROTEIN; ALBUMIN; SIGNAL.
FT SIGNAL 1 21
FT PROPEP 22 37
FT CHAIN 38 72
FT PROPEP 73 81
SMALL SUBUNIT (BY SIMILARITY).

FT CHAIN 82 162 LARGE SUBUNIT (BY SIMILARITY).
SQ SEQUENCE 164 AA; 18762 MW; D4F9F482 CRC32;

Query Match 52.6%; Score 40; DB 1; Length 164;
Best Local Similarity 46.2%; Pred. NO. 4.9;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 COSKSSEIRCGK 13
DB 58 CQWMSKQMRGR 70

RESULT 13
CP18 RABIT STANDARD; PRT; 171 AA.
ID CP18 RABIT STANDARD; PRT; 171 AA.
AC P25230;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE ANTIMICROBIAL PROTEIN CAP18 PRECURSOR (18 KD LIPOPOLYSACCHARIDE-
DE BINDING PROTEIN) (18 KD CATIONIC PROTEIN) (CAP18-A).
GN CAP18
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
[1]
RN SEQUENCE FROM N.A.; AND SEQUENCE OF 135-159.
RP TISSUE-BONE MARROW;
RC MEDLINE; 91354246;
RA LARRICK J.W.; MORGAN J.G.; PALINGS I.; HIRATA M.; YEN M.H.;
RT "Complementary DNA sequence of rabbit CAP18 -- a unique
lipopolysaccharide binding protein.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 179:170-175(1991).
[2]
RN SEQUENCE OF 135-159, AND CHARACTERIZATION.
RP MEDLINE; 94178952.
RA HIRATA M.; SHIMOMURA Y.; YOSHIDA M.; MORGAN J.G.; PALINGS I.;
RA WILSON D.; YEN M.H.; WRIGHT S.C.; LARRICK J.W.;
RT "Characterization of a rabbit cationic protein (CAP18) with
lipopolysaccharide-inhibitory activity.";
RL INFECT. IMMUN. 62:1421-1426(1994).
[3]
RN SEQUENCE OF 135-154, AND CHARACTERIZATION.
RP MEDLINE; 94075827.
RA LARRICK J.W.; HIRATA M.; ZHENG H.; ZHONG J.; BOLIN D.;
RA CAVAILLON J.-M.; WARREN H.S.; WRIGHT S.C.;
RT "A novel granulocyte-derived peptide with lipopolysaccharide-
neutralizing activity.";
RL J. IMMUNOL. 152:231-240(1994).
[4]
RN CHARACTERIZATION.
RP MEDLINE; 94148064.
RA TOSSI A.; SOCCCHI M.; SKERLAVAJ B.; GENNARO R.;
RT "Identification and characterization of a primary antibacterial
domain in CAP18, a lipopolysaccharide binding protein from rabbit
leukocytes.";
RL FEBS LETT. 339:108-112(1994).
[5]
RN STRUCTURE BY NMR OF 135-166.
RP MEDLINE; 95377455.
RA CHEN C.; BROCK R.; LUH F.; CHOU P.-J.; LARRICK J.W.; HUANG R.-F.;
RA HUANG T.-H.;
RT "The solution structure of the active domain of CAP18 -- a
lipopolysaccharide binding protein from rabbit leukocytes.";
RL FEBS LETT. 370:46-52(1995).
CC -!- FUNCTION: CAP18 BINDS TO THE LIPID A MOIETY OF BACTERIAL
CC LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER
CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. HAS ANTIBIOTIC ACTIVITY.
CC -!- TISSUE SPECIFICITY: NEUTROPHILS.
CC -!- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



DE NONSTRUCTURAL PROTEIN NS1 (FRAGMENT).  
 OS DENGUE VIRUS TYPE 1 (STRAIN 836-1).  
 OC VIRUSES: SSRA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;  
 FLAVIVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89293078.  
 RA CHU M.C., O'ROURKE E.J., TRENT D.W.;  
 RT "Genetic relatedness among structural protein genes of dengue 1 virus  
 strains.";  
 RL J. GEN. VIROL. 70:1701-1712(1989).  
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF THREE PROTEINS:  
 CC PROTEIN PM, PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A  
 CC COMPLEX OF PROTEIN C AND MRNA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D00503; D1000849; -  
 DR PIR: C32401; C32401.  
 DR PFAM: PF00869; Flavi\_glycoprot; 1.  
 DR PFAM: PF01003; Flavi\_capsid; 1.  
 DR PFAM: PF01004; Flavi\_M; 1.  
 DR HSP: P14336; 1SVB.  
 KW POLYPROTEIN: GLYCOPROTEIN; CORE PROTEIN; COAT PROTEIN;  
 KW ENVELOPE PROTEIN; TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.  
 FT CHAIN 1 114 CAPSID PROTEIN C.  
 FT PROPEP 115 205 ENVELOPE GLYCOPROTEIN M.  
 FT CHAIN 206 280 MAJOR ENVELOPE PROTEIN E.  
 FT CHAIN 281 775 NONSTRUCTURAL PROTEIN NS1.  
 FT CHAIN 776 >792 POTENTIAL.  
 FT TRANSEM 46 67 POTENTIAL.  
 FT TRANSEM 102 118 POTENTIAL.  
 FT TRANSEM 244 263 POTENTIAL.  
 FT TRANSEM 266 281 POTENTIAL.  
 FT TRANSEM 715 735 POTENTIAL.  
 FT TRANSEM 755 773 POTENTIAL.  
 FT DISULFID 283 310 BY SIMILARITY.  
 FT DISULFID 340 396 BY SIMILARITY.  
 FT DISULFID 354 385 BY SIMILARITY.  
 FT DISULFID 372 401 BY SIMILARITY.  
 FT DISULFID 465 565 BY SIMILARITY.  
 FT DISULFID 582 613 BY SIMILARITY.  
 FT CARBOHYD 183 183 POTENTIAL.  
 FT CARBOHYD 347 347 POTENTIAL.  
 FT CARBOHYD 433 433 POTENTIAL.  
 FT NON\_TER 792 792  
 SQ SEQUENCE 792 AA; 86913 MW; 472FC6F4 CRC32;

Query Match 53.98; Score 41; DB 1; Length 792;  
 Best Local Similarity 50.08; Pred. No. 15;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 :||| :|||  
 Db 782 NWKGLKCG 791

RESULT 9  
 POLG\_DENIA STANDARD; PRT; 792 AA.  
 AC P27912;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX

DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;  
 OS NONSTRUCTURAL PROTEIN NS1 (FRAGMENT).  
 OS DENGUE VIRUS TYPE 1 (STRAIN AHF 82-80).  
 OC VIRUSES: SSRA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;  
 FLAVIVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89293078.  
 RA CHU M.C., O'ROURKE E.J., TRENT D.W.;  
 RT "Genetic relatedness among structural protein genes of dengue 1 virus  
 strains.";  
 RL J. GEN. VIROL. 70:1701-1712(1989).  
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF THREE PROTEINS:  
 CC PROTEIN PM, PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A  
 CC COMPLEX OF PROTEIN C AND MRNA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D00502; D1000848; -  
 DR PIR: B32401; B32401.  
 DR PFAM: PF00869; Flavi\_glycoprot; 1.  
 DR PFAM: PF01003; Flavi\_capsid; 1.  
 DR PFAM: PF01004; Flavi\_M; 1.  
 DR HSP: P14336; 1SVB.  
 KW POLYPROTEIN: GLYCOPROTEIN; CORE PROTEIN; COAT PROTEIN;  
 KW ENVELOPE PROTEIN; TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.  
 FT CHAIN 1 114 CAPSID PROTEIN C.  
 FT PROPEP 115 205 ENVELOPE GLYCOPROTEIN M.  
 FT CHAIN 206 280 MAJOR ENVELOPE PROTEIN E.  
 FT CHAIN 281 775 NONSTRUCTURAL PROTEIN NS1.  
 FT CHAIN 776 >792 POTENTIAL.  
 FT TRANSEM 46 67 POTENTIAL.  
 FT TRANSEM 102 118 POTENTIAL.  
 FT TRANSEM 244 263 POTENTIAL.  
 FT TRANSEM 266 281 POTENTIAL.  
 FT TRANSEM 715 735 POTENTIAL.  
 FT TRANSEM 755 773 POTENTIAL.  
 FT DISULFID 283 310 BY SIMILARITY.  
 FT DISULFID 340 396 BY SIMILARITY.  
 FT DISULFID 354 385 BY SIMILARITY.  
 FT DISULFID 372 401 BY SIMILARITY.  
 FT DISULFID 465 565 BY SIMILARITY.  
 FT DISULFID 582 613 BY SIMILARITY.  
 FT CARBOHYD 183 183 POTENTIAL.  
 FT CARBOHYD 347 347 POTENTIAL.  
 FT CARBOHYD 433 433 POTENTIAL.  
 FT NON\_TER 792 792  
 SQ SEQUENCE 792 AA; 86884 MW; 493CD491 CRC32;

Query Match 53.98; Score 41; DB 1; Length 792;  
 Best Local Similarity 50.08; Pred. No. 15;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 :||| :|||  
 Db 782 NWKGLKCG 791

RESULT 10  
 POLG\_DENIC STANDARD; PRT; 791 AA.  
 AC P27913;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)



```
DR PFAM; PF00948; Flavi_NS1; 1.
DR PFAM; PF00949; Flavi_helicase; 1.
DR PFAM; PF00972; Flavi_NS5; 1.
DR PFAM; PF01002; Flavi_NS2B; 1.
DR PFAM; PF01003; Flavi_capsid; 1.
DR PFAM; PF01004; Flavi_M; 1.
DR PFAM; PF01005; Flavi_NS2A; 1.
DR HSP; P14336; 1SVB.
KW POLYPROTEIN; GLYCOPROTEIN; TRANSFERASE; RNA-DIRECTED RNA POLYMERASE;
KW CORE PROTEIN; COAT PROTEIN; ENVELOPE PROTEIN; HELICASE; ATP-BINDING;
KW TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.
FT CHAIN 1 114
FT PROPEP 115 205
FT CHAIN 206 280
FT CHAIN 281 775
FT CHAIN 776 1127
FT CHAIN 1128 1345
FT CHAIN 1346 1475
FT CHAIN 1476 2090
FT CHAIN 2091 2376
FT CHAIN 2377 2488
FT CHAIN 2489 3388
FT NP_BIND 1668 1675
FT SITE 1759 1762
FT TRANSMEM 43 59
FT TRANSMEM 101 117
FT TRANSMEM 268 284
FT TRANSMEM 727 743
FT TRANSMEM 757 773
FT DISULFID 283 310
FT DISULFID 340 396
FT DISULFID 354 385
FT DISULFID 372 401
FT DISULFID 465 565
FT DISULFID 582 613
FT CARBOHYD 183 183
FT CARBOHYD 347 347
FT CARBOHYD 433 433
FT CARBOHYD 905 905
FT CARBOHYD 982 982
FT CARBOHYD 1134 1134
FT CARBOHYD 1174 1174
FT CARBOHYD 2298 2298
FT CARBOHYD 2302 2302
FT CARBOHYD 2454 2454
FT CARBOHYD 2482 2482
SQ SEQUENCE 3388 AA; 379213 MW; 8D0A2D62 CRC32;

Query Match 59.2%; Score 45; DB 1; Length 3388;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 3 SWKSSEIRCG 12
Db 782 SWKNKELKCG 791
||||:|:|:|

RESULT 7
ID POLG_DEN2T STANDARD; PRT; 1683 AA.
AC P27914;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DE GENOME POLYPROTEIN [CONTAINS: MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL
DE PROTEINS NS1 AND NS2A; HELICASE (NS3)] [FRAGMENTS].
OS DENGUE VIRUS TYPE 2 (STRAIN TONGA 1974).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;
OC FLAVIVIRUS.
RN [1]
RP SEQUENCE OF 1-495 FROM N.A.
RX MEDLINE; 91016939.

CHEN W., MAGUIRE T.;
"Nucleotide sequence of the envelope glycoprotein gene of a dengue-2
virus isolated during an epidemic of benign dengue fever in Tonga in
1974.";
RL NUCLEIC ACIDS RES. 18:5889-5889(1990).
[2]
RP SEQUENCE OF 496-1683 FROM N.A.
OU X., CHEN W., MAGUIRE T.;
SUBMITTED (MAR-1991) TO EMBL/GENBANK/DBJ DATA BANKS.
RL SUBMITT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X54319; G59304; -
DR EMBL; X57469; G930021; -
DR EMBL; X57468; G930022; -
DR PFAM; PF00869; Flavi_glycoprot; 1.
DR PFAM; PF00948; Flavi_NS1; 1.
DR PFAM; PF00949; Flavi_helicase; 1.
DR PFAM; PF01005; Flavi_NS2A; 1.
DR HSP; P14336; 1SVB.
KW POLYPROTEIN; GLYCOPROTEIN; CORE PROTEIN; COAT PROTEIN;
ENVELOPE PROTEIN; TRANSMEMBRANE; HELICASE; ATP-BINDING;
KW NONSTRUCTURAL PROTEIN.
FT CHAIN 1 495
FT CHAIN 496 847
FT CHAIN 848 1065
FT NON_CONS 1065 1066
FT CHAIN 1066 1683
FT NP_BIND 1258 1265
FT SITE 1349 1352
FT TRANSMEM 457 470
FT TRANSMEM 472 495
FT DISULFID 3 30
FT DISULFID 60 116
FT DISULFID 74 105
FT DISULFID 92 121
FT DISULFID 185 285
FT DISULFID 302 333
FT CARBOHYD 67 67
FT CARBOHYD 153 153
FT NON_TER 1683 1683
SQ SEQUENCE 1683 AA; 187439 MW; C3F3A88D CRC32;

Query Match 59.2%; Score 45; DB 1; Length 1683;
Best Local Similarity 60.0%; Pred. No. 6.9;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 3 SWKSSEIRCG 12
Db 502 SWKNKELKCG 511
||||:|:|:|

RESULT 8
ID POLG_DEN18 STANDARD; PRT; 792 AA.
AC P27910;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX
DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;
```

CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:  
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  
CC PROTEIN C AND MRNA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M29095; G323448; -  
DR PIR: A28646; A28646.  
DR PFAM: PF00869; Flavi\_glycoprot; 1.  
DR PFAM: PF00948; Flavi\_Ns1; 1.  
DR PFAM: PF00949; Flavi\_helicase; 1.  
DR PFAM: PF00972; Flavi\_Ns5; 1.  
DR PFAM: PF01002; Flavi\_Ns2B; 1.  
DR PFAM: PF01003; Flavi\_capsid; 1.  
DR PFAM: PF01004; Flavi\_M; 1.  
DR PFAM: PF01005; Flavi\_Ns2A; 1.  
DR HSP: P14336; ISVB.  
DR POLYPROTEIN; GLYCOPROTEIN; TRANSFERASE; RNA-DIRECTED RNA POLYMERASE;  
KW CORE PROTEIN; COAT PROTEIN; ENVELOPE PROTEIN; HELICASE; ATP-BINDING;  
KW TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.  
FT CHAIN 1 114  
FT PROPEP 115 205  
FT CHAIN 206 280  
FT CHAIN 281 775  
FT CHAIN 776 1127  
FT CHAIN 1128 1345  
FT CHAIN 1346 1475  
FT CHAIN 1476 2093  
FT CHAIN 2094 2243  
FT CHAIN 2244 2491  
FT CHAIN 2492 3391  
FT NP\_BIND 1668 1675  
FT SITE 1759 1762  
FT TRANSMEM 50 66  
FT TRANSMEM 102 118  
FT TRANSMEM 268 284  
FT TRANSMEM 727 743  
FT TRANSMEM 757 773  
FT TRANSMEM 1158 1174  
FT TRANSMEM 1272 1288  
FT TRANSMEM 1294 1310  
FT TRANSMEM 1351 1367  
FT TRANSMEM 1373 1389  
FT TRANSMEM 1448 1464  
FT TRANSMEM 2148 2164  
FT TRANSMEM 2174 2190  
FT TRANSMEM 2197 2213  
FT TRANSMEM 2227 2243  
FT TRANSMEM 2352 2368  
FT TRANSMEM 2411 2427  
FT DISULFID 283 310  
FT DISULFID 340 396  
FT DISULFID 354 385  
FT DISULFID 372 401  
FT DISULFID 465 565  
FT DISULFID 582 613  
FT CARBOHYD 183 183  
FT CARBOHYD 347 347  
FT CARBOHYD 433 433  
FT CARBOHYD 905 905  
FT CARBOHYD 982 982  
FT CARBOHYD 1134 1134  
FT CARBOHYD 2301 2301  
FT CARBOHYD 2305 2305  
FT CARBOHYD 2457 2457  
FT CARBOHYD 2485 2485

FT CARBOHYD 2665 2665 POTENTIAL.  
FT CARBOHYD 2704 2704 POTENTIAL.  
FT CARBOHYD 2714 2714 POTENTIAL.  
SQ SEQUENCE 3391 AA; 379494 MW; 4C294F49 CRC32;  
  
Query Match 59.2%; Score 45; DB 1; Length 3391;  
Best Local Similarity 60.0%; Pred. No. 14;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 SWKSSIRCG 12  
DB 782 SWKNKELKCG 791  
|||:|:|:  
  
RESULT 6  
POLG\_DEN2P STANDARD; PRT: 3388 AA.  
AC P12823; Q88646; Q88647; Q88648; Q88649; Q88650; Q88651; Q88652;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX  
DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;  
DE NONSTRUCTURAL PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3);  
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (NS5)].  
OS DENGUE VIRUS TYPE 2 (STRAIN PR159/S1).  
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;  
OC FLAVIVIRUS.  
CC [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE; 88101365.  
RA HAHN Y.S., GALLER R., HUNKAPILLER T., DALRYMPLE J.M., STRAUSS J.H.,  
RA STRAUSS E.G.;  
RA "Nucleotide sequence of dengue 2 RNA and comparison of the encoded  
RT proteins with those of other flaviviruses.";  
RL VIROLOGY 162:167-180(1988).  
RN [2]  
RP C-TERMINAL OF M, E, AND NS1.  
RC STRAIN-NEW-GUINEA;  
RX MEDLINE; 89299482.  
RA WRIGHT P.J., CAUCHI M.R., NG M.L.;  
RT "Definition of the carboxy termini of the three glycoproteins  
RT specified by dengue virus type 2.";  
RL VIROLOGY 171:61-67(1989).  
CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE  
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.  
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.  
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:  
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  
CC PROTEIN C AND MRNA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M19197; G323655; -  
DR EMBL: M19197; E12056; ALT\_TERM.  
DR EMBL: M19197; E12057; ALT\_SEQ.  
DR EMBL: M19197; E12058; ALT\_SEQ.  
DR EMBL: M19197; E12059; ALT\_SEQ.  
DR EMBL: M19197; E12060; ALT\_SEQ.  
DR EMBL: M19197; E12061; ALT\_SEQ.  
DR EMBL: M19197; E12062; ALT\_SEQ.  
DR EMBL: M19197; E12063; ALT\_SEQ.  
DR EMBL: M19197; E12064; ALT\_SEQ.  
DR PIR: A29972; GNNVDP.  
DR PFAM: PF00869; Flavi\_glycoprot; 1.

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  
PROTEIN C AND MRNA.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M20558; G323450; -;  
EMBL; M15075; G323651; -;  
EMBL; M14969; G807601; -;  
EMBL; M14970; -; NOT\_ANNOTATED\_CDS.

PIR; A29199; GNVVJA.

PFAM; PF00869; Flavi\_glycoprot; 1.

PFAM; PF00948; Flavi\_NS1; 1.

PFAM; PF00949; Flavi\_helicase; 1.

PFAM; PF00972; Flavi\_NS5; 1.

PFAM; PF01002; Flavi\_NS2B; 1.

PFAM; PF01003; Flavi\_capsid; 1.

PFAM; PF01004; Flavi\_M; 1.

PFAM; PF01005; Flavi\_NS2A; 1.

HSP; P14336; 1SVB.

POLYPEPTIDE; GLYCOPROTEIN; TRANSFERASE; RNA-DIRECTED RNA POLYMERASE;  
CORE PROTEIN; COAT PROTEIN; ENVELOPE PROTEIN; HELICASE; ATP-BINDING;  
TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.

CHAIN 1 114

PROPEP 115 205

CHAIN 206 280

CHAIN 281 775

CHAIN 776 1127

CHAIN 1128 1345

CHAIN 1346 1475

CHAIN 1476 2093

CHAIN 2094 2243

CHAIN 2244 2491

CHAIN 2492 3391

NP\_BIND 1668 1675

SITE 1759 1762

TRANSMEM 50 66

TRANSMEM 102 118

TRANSMEM 268 284

TRANSMEM 727 743

TRANSMEM 757 773

TRANSMEM 1158 1174

TRANSMEM 1272 1288

TRANSMEM 1294 1310

TRANSMEM 1351 1367

TRANSMEM 1373 1389

TRANSMEM 1448 1464

TRANSMEM 2148 2164

TRANSMEM 2174 2190

TRANSMEM 2197 2213

TRANSMEM 2227 2243

TRANSMEM 2352 2368

TRANSMEM 2411 2427

TRANSMEM 2983 310

TRANSMEM 340 396

TRANSMEM 354 385

TRANSMEM 372 401

TRANSMEM 465 565

TRANSMEM 582 613

TRANSMEM 183 183

TRANSMEM 347 347

TRANSMEM 433 433

TRANSMEM 905 905

TRANSMEM 982 982

TRANSMEM 1134 1134

TRANSMEM 2301 2301

TRANSMEM 2305 2305

TRANSMEM 2457 2457

FT CARBOHYD 2485 2485 POTENTIAL.  
FT CONFLICT 1044 1044 H -> D (IN REF. 3).  
FT CONFLICT 1231 1231 A -> T (IN REF. 3).  
FT CONFLICT 1236 1236 A -> V (IN REF. 3).  
FT CONFLICT 1263 1263 I -> M (IN REF. 3).  
FT CONFLICT 1266 1266 N -> K (IN REF. 3).  
FT CONFLICT 1301 1301 L -> V (IN REF. 3).  
FT CONFLICT 1308 1308 L -> F (IN REF. 3).  
FT CONFLICT 1342 1342 S -> N (IN REF. 3).  
FT CONFLICT 1454 1454 V -> L (IN REF. 3).  
FT CONFLICT 1503 1503 R -> K (IN REF. 3).  
FT CONFLICT 1552 1552 ISY -> VSC (IN REF. 3).  
FT CONFLICT 1595 1595 T -> A (IN REF. 3).  
FT CONFLICT 1615 1615 VDR -> IDK (IN REF. 3).  
FT CONFLICT 1661 1661 KR -> RK (IN REF. 3).  
FT CONFLICT 1685 1685 E -> G (IN REF. 3).  
FT CONFLICT 1820 1820 N -> S (IN REF. 3).  
FT CONFLICT 1846 1846 I -> T (IN REF. 3).  
FT CONFLICT 1858 1858 I -> T (IN REF. 3).  
FT CONFLICT 1874 1874 A -> T (IN REF. 3).  
FT CONFLICT 1878 1878 D -> N (IN REF. 3).  
FT CONFLICT 2169 2169 K -> R (IN REF. 3).  
FT CONFLICT 2265 2265 ES -> OP (IN REF. 3).  
FT CONFLICT 2330 2330 IH -> MD (IN REF. 3).  
FT CONFLICT 2355 2355 L -> F (IN REF. 3).  
FT CONFLICT 2391 2391 DG -> GR (IN REF. 3).  
FT CONFLICT 2418 2418 I -> V (IN REF. 3).  
FT CONFLICT 2647 2647 I -> V (IN REF. 3).  
FT CONFLICT 2687 2687 T -> A (IN REF. 3).  
SQ SEQUENCE 3391 AA; 379244 MW; D250C40B CRC32;

Query Match 59.2%; Score 45; DB 1; Length 3391;  
Best Local Similarity 60.0%; Pred. No. 14;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 3 SWKSSEIRCG 12

Db 782 SWKNKELKCG 791

RESULT 5

POLG\_DEN2N ID POLG\_DEN2N STANDARD; PRT; 3391 AA.

AC P14340;

DT 01-JAN-1990 (REL. 13, CREATED)

DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE GENOME POLYPEPTIDE [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX

DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;

DE NONSTRUCTURAL PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3);

DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (NS5)].

DE DENGUE VIRUS TYPE 2 (STRAIN NEW GUINEA C).

OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;

OC FLAVIVIRUS.

RN [1]

RN SEQUENCE FROM N.A.

RX MEDLINE; 89232751.

RA IRIE K., MOHAN P.M., SASAGURI Y., PUTNAK R., PADMANABHAN R.;

RT "Sequence analysis of cloned dengue virus type 2 genome (New Guinea-C

RL strain).";

RL GENE 75:197-211(1989).

RN [2]

RN SEQUENCE OF 749-1225 FROM N.A.

RX MEDLINE; 88160069.

RA PUTNAK J.R., CHARLES P.C., PADMANABHAN R., IRIE K., HOKE C.H.,

RA BURKE D.S.;

RT "Functional and antigenic domains of the dengue-2 virus nonstructural

RL glycoprotein NS-1.,";

RL VIROLOGY 163:93-103(1988).

CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE

CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

FT CARBOHYD 2305 2305 POTENTIAL.  
 FT CARBOHYD 2457 2457 POTENTIAL.  
 FT CARBOHYD 2485 2485 POTENTIAL.  
 FT CARBOHYD 2665 2665 POTENTIAL.  
 FT CARBOHYD 2704 2704 POTENTIAL.  
 FT CARBOHYD 2714 2714 POTENTIAL.  
 SQ SEQUENCE 3391 AA; 379878 MW; 70570314 CRC32;

Query Match 59.2%; Score 45; DB 1; Length 3391;  
 Best Local Similarity 60.0%; Pred. No. 14;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 |||:|:|:|

Db 782 SWKNKELKCG 791

RESULT 3  
 POLG\_DEN2D STANDARD; PRT; 1127 AA.  
 AC P30026; Q66450;  
 DT 01-APR-1993 (REL. 25, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX  
 DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;  
 DE NONSTRUCTURAL PROTEIN NS1] (FRAGMENT).  
 OS DENGUE VIRUS TYPE 2 (STRAIN D2-04).  
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;  
 RN FLAVIVIRUS.  
 [1]  
 RP SEQUENCE OF 1-775 FROM N.A.  
 RA YANG P.Y., LAM S.K.;  
 RT "The nucleotide and encoded amino acid sequences of the structural  
 RT protein gene of D2-04 virus strain isolated in China."  
 RL CHIN. J. MICROBIOL. IMMUNOL. 11:341-344(1991).  
 RN [2]  
 RP SEQUENCE OF 776-1127 FROM N.A.  
 RA YANG P.Y., KAUTNER I.M., KOH C.L., LAM S.K.;  
 RT "Nucleotide and encoded amino acid sequences of the nonstructural  
 RT protein NS1 gene of a dengue-2 virus isolated in China."  
 RL CHIN. J. MICROBIOL. IMMUNOL. 11:9-12(1991).  
 CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:  
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF  
 CC PROTEIN C AND MRNA.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: U1978; G847734; -  
 DR PIR: JC1007; GNWVD2.  
 DR PFAM: PF00869; Flavi\_glycoprot; 1.  
 DR PFAM: PF00948; Flavi\_Ns1; 1.  
 DR PFAM: PF01003; Flavi\_capsid; 1.  
 DR PFAM: PF01004; Flavi\_M; 1.  
 DR HSP: P14336; 1SVB.  
 KW POLYPROTEIN; GLYCOPROTEIN; CORE PROTEIN; COAT PROTEIN;  
 KW ENVELOPE PROTEIN; TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.  
 FT CHAIN 1 114  
 FT PROPEP 115 205  
 FT CHAIN 206 280  
 FT CHAIN 281 775  
 FT CHAIN 776 1127  
 FT CHAIN 101 117  
 FT TRANSMEM 727 743  
 FT TRANSMEM 757 773

FT DISULFID 283 310 BY SIMILARITY.  
 FT DISULFID 340 396 BY SIMILARITY.  
 FT DISULFID 354 385 BY SIMILARITY.  
 FT DISULFID 372 401 BY SIMILARITY.  
 FT DISULFID 465 565 BY SIMILARITY.  
 FT DISULFID 582 613 BY SIMILARITY.  
 FT CARBOHYD 183 183 POTENTIAL.  
 FT CARBOHYD 347 347 POTENTIAL.  
 FT CARBOHYD 433 433 POTENTIAL.  
 FT CARBOHYD 905 905 POTENTIAL.  
 FT CARBOHYD 982 982 POTENTIAL.  
 FT NON\_TER 1127 1127  
 SQ SEQUENCE 1127 AA; 125835 MW; A0633DBC CRC32;

Query Match 59.2%; Score 45; DB 1; Length 1127;  
 Best Local Similarity 60.0%; Pred. No. 4.7;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 |||:|:|:|

Db 782 SWKNKELKCG 791

RESULT 4  
 POLG\_DEN2J STANDARD; PRT; 3391 AA.  
 AC P07564; P07565;  
 DT 01-APR-1988 (REL. 07, CREATED)  
 DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX  
 DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;  
 DE NONSTRUCTURAL PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3);  
 DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (NS5)].  
 OS DENGUE VIRUS TYPE 2 (STRAIN JAMAICA).  
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;  
 RN FLAVIVIRUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 86265864.  
 RA DEUBEL V., KINNEY R.M., TRENT D.W.;  
 RT "Nucleotide sequence and deduced amino acid sequence of the  
 RT nonstructural proteins of dengue type 2 virus, Jamaica genotype:  
 RT comparative analysis of the full-length genome."  
 RL VIROLOGY 165:234-244(1988).  
 RN [2]  
 RP SEQUENCE OF 1-791 FROM N.A.  
 RX MEDLINE; 87071658.  
 RA DEUBEL V., KINNEY R.M., TRENT D.W.;  
 RT "Nucleotide sequence and deduced amino acid sequence of the  
 RT structural proteins of dengue type 2 virus, Jamaica genotype."  
 RL VIROLOGY 155:365-377(1986).  
 RN [3]  
 RP SEQUENCE OF 1044-1928 AND 2119-2761 FROM N.A.  
 RX MEDLINE; 87106861.  
 RA YAGASHI T., VAKHARIA V.N., PAGE K., SASAGURI Y., FEIGHNY R.,  
 RA PADMANABHAN R.;  
 RT "Partial sequence analysis of cloned dengue virus type 2 genome."  
 RL GENE 46:257-267(1986).  
 RN [4]  
 RP C-TERMINAL OF M, E, AND NS1.  
 RC STRAIN-NEW-GUINEA.  
 RX MEDLINE; 89299482.  
 RA WRIGHT P.J., CAUCHI M.R., NG M.L.;  
 RT "Definition of the carboxy termini of the three glycoproteins  
 RT specified by dengue virus type 2."  
 RL VIROLOGY 171:61-67(1989).  
 CC -!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE  
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.  
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.  
 CC -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A  
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:09 ; Search time 71.09 Seconds  
(without alignments)  
5.169 Million cell updates/sec

Title: US-09-124-280A-29  
Perfect score: 76  
Sequence: 1 CQSWKSSEIRCGK 13

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	59.2	3391	1	POLG_DEN26
2	45	59.1	3391	1	POLG_DEN27
3	45	59.2	1127	1	POLG_DEN2D
4	45	59.2	3391	1	POLG_DEN2J
5	45	59.2	3391	1	POLG_DEN2N
6	45	59.2	3388	1	POLG_DEN2P
7	45	59.2	1683	1	POLG_DEN2T
8	41	53.9	792	1	POLG_DEN18
9	41	53.9	792	1	POLG_DEN1A
10	41	53.9	791	1	POLG_DEN1C
11	41	53.9	3390	1	POLG_DEN3
12	40	52.6	184	1	2SS3-ARATH
13	40	52.6	171	1	CP18-RABIT
14	40	52.6	3396	1	POLG_DEN1S
15	40	52.6	1226	1	POLG_DEN1W
16	39	51.3	528	1	HS60-AJECA
17	39	51.3	414	1	LIP5-YEAST
18	39	51.3	106	1	LV4E-HUMAN
19	39	51.3	122	1	PA21-AGKHA
20	39	51.3	3386	1	POLG_DEN4
21	39	51.3	364	1	PVH1-YEAST
22	38.5	50.7	695	1	CASP-MESAU
23	38	50.0	389	1	SH1B-CAVPO
24	38	50.0	386	1	SH1B-CRIGR
25	38	50.0	388	1	SH1B-DIDMA
26	38	50.0	390	1	SH1B-HUMAN
27	38	50.0	386	1	SH1B-MOUSE
28	38	50.0	390	1	SH1B-RABIT
29	38	50.0	386	1	SH1B-RAT
30	38	50.0	386	1	SH1B-SPAEH
31	38	50.0	377	1	SH1D-CANFA
32	38	50.0	376	1	SH1D-CAVPO
33	38	50.0	377	1	SH1D-HUMAN
34	38	50.0	374	1	SH1D-MOUSE
35	38	50.0	291	1	SH1D-PIG
36	38	50.0	377	1	SH1D-RABIT
37	38	50.0	374	1	SH1D-RAT
38	38	50.0	466	1	5HT-HELVI
39	38	50.0	437	1	AFLR-ASPL
40	38	50.0	444	1	AFLR-ASPPA
41	38	50.0	329	1	ESG2-TRYBB
42	38	50.0	330	1	MTHT-METIF
43	38	50.0	284	1	PURU-SYNY3

44	38	50.0	471	1	YHA2-CRYPA	P10942 cryphonectr
45	37.5	49.3	1468	1	DPOA-YEAST	P13382 saccharomyc
ALIGNMENTS						
RESULT 1						
ID	POLG_DEN26	STANDARD;	PRT;	3391	AA.	
AC	P29990:					
DT	01-APR-1993	(REL. 25, CREATED)				
DT	01-APR-1993	(REL. 25, LAST SEQUENCE UPDATE)				
DT	15-DEC-1998	(REL. 37, LAST ANNOTATION UPDATE)				
DE	GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL PROTEINS NS1, NS2A, NS2B, NS4A AND NS4B; HELICASE (NS3); RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (NS5)].					
OS	DENGUE VIRUS TYPE 2 (STRAIN 16681).					
OC	VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE; FLAVIVIRUS.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE;	92188532.				
RA	BLOK J., MCWILLIAM S.M., BUTLER H.C., GIBBS A.J., WEILLER G., HERRING B.L., HEMSLEY A.C., AASKOV J.G., YOKSAN S., BHAMARAPRATI N.;					
RT	"Comparison of a dengue-2 virus and its candidate vaccine derivative: sequence relationships with the flaviviruses and other viruses.";					
RL	VIROLOGY 197:573-590(1992).					
CC	-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.					
CC	-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.					
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
CC	EMBL;	M84727;	G323473;			
DR	PIR;	A42451;	GNV16.			
DR	PFAM;	PF00869;	Flavi_glycoprot;	1.		
DR	PFAM;	PF00948;	Flavi_NS1;	1.		
DR	PFAM;	PF00949;	Flavi_helicase;	1.		
DR	PFAM;	PF00972;	Flavi_NS5;	1.		
DR	PFAM;	PF01002;	Flavi_NS2B;	1.		
DR	PFAM;	PF01003;	Flavi_capsid;	1.		
DR	PFAM;	PF01004;	Flavi_M;	1.		
DR	PFAM;	PF01005;	Flavi_NS2A;	1.		
DR	HSSP;	P14336;	1SVB.			
KW	POLYPROTEIN; GLYCOPROTEIN; TRANSFERASE; RNA-DIRECTED RNA POLYMERASE; CORE PROTEIN; COAT PROTEIN; ENVELOPE PROTEIN; HELICASE; ATP-BINDING; TRANSMEMBRANE; NONSTRUCTURAL PROTEIN.					
KW	CHAIN	1	114			
FT	PROPEP	115	205			
FT	CHAIN	206	280			
FT	CHAIN	281	775			
FT	CHAIN	776	1127			
FT	CHAIN	1128	1345			
FT	CHAIN	1346	1474			
FT	CHAIN	1475	2093			
FT	CHAIN	2094	2243			
FT	CHAIN	2244	2491			
FT	CHAIN	2492	3391			
FT	NP_BIND	1668	1675			
FT	SITE	1759	1762			

**THIS PAGE BLANK (USPTO)**

D43763  
nonstructural protein NS1 - dengue virus type 2 (strain D80-100)  
C:Species: dengue virus type 2  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 29-May-1998  
C:Accession: D43763  
R:Block, J.; Gibbs, A.J.; McWilliam, S.M.; Vitarana, U.T.  
Arch. Virol. 118, 209-223, 1991  
A:Title: NS1 gene sequences from eight dengue-2 viruses and their evolutionary relations  
A:Reference number: A43763  
A:Accession: D43763  
A:Molecule type: genomic RNA  
A:Residues: 1-352 <BLO>  
A:Cross-references: GB:M58489; NID:g323457; PID:g323458  
C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology  
C:Keywords: nonstructural protein

Query Match 59.2%; Score 45; DB 2; Length 352;  
Best Local Similarity 60.0%; Pred. No. 2.6;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
|||:|::||  
Db 7 SWKNKELKCG 16

RESULT 14  
E43763  
nonstructural protein NS1 - dengue virus type 2 (strain D80-141)  
C:Species: dengue virus type 2  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 29-May-1998  
C:Accession: E43763  
R:Block, J.; Gibbs, A.J.; McWilliam, S.M.; Vitarana, U.T.  
Arch. Virol. 118, 209-223, 1991  
A:Title: NS1 gene sequences from eight dengue-2 viruses and their evolutionary relations  
A:Reference number: A43763  
A:Accession: E43763  
A:Molecule type: genomic RNA  
A:Residues: 1-352 <BLO>  
A:Cross-references: GB:M58490; NID:g323459; PID:g323460  
C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology  
C:Keywords: nonstructural protein

Query Match 59.2%; Score 45; DB 2; Length 352;  
Best Local Similarity 60.0%; Pred. No. 2.6;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
|||:|::||  
Db 7 SWKNKELKCG 16

RESULT 15  
F43763  
nonstructural protein NS1 - dengue virus type 2 (strain SL767)  
C:Species: dengue virus type 2  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 29-May-1998  
C:Accession: F43763  
R:Block, J.; Gibbs, A.J.; McWilliam, S.M.; Vitarana, U.T.  
Arch. Virol. 118, 209-223, 1991  
A:Title: NS1 gene sequences from eight dengue-2 viruses and their evolutionary relations  
A:Reference number: A43763  
A:Accession: F43763  
A:Molecule type: genomic RNA  
A:Residues: 1-352 <BLO>  
A:Cross-references: GB:M58491; NID:g323461; PID:g323462  
C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology  
C:Keywords: nonstructural protein

Query Match 59.2%; Score 45; DB 2; Length 352;  
Best Local Similarity 60.0%; Pred. No. 2.6;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
|||:|::||  
Db 7 SWKNKELKCG 16

Search completed: September 7, 1999, 23:22:44  
Job time: 951 sec



A:Description: Nucleotide and encoded amino acid sequences of the nonstructural protein  
 A:Reference number: S34368  
 A:Accession: S34368  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-352 <SHI>  
 A:Cross-references: EMBL:X69191; NID:g312759; PID:g1334832  
 C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology

Query Match 59.2%; Score 45; DB 2; Length 352;  
 Best Local Similarity 60.0%; Pred. No. 2.6;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 |||:|:|:|  
 Db 7 SWKNKELKCG 16

RESULT 9  
 PQ0507  
 polyprotein - dengue virus type 2 (strains Tonga 251 and Tonga 1974) (fragment)  
 N:Contains: nonstructural protein NS1; nonstructural protein NS2a; nonstructural protein  
 C:Species: dengue virus type 2  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 29-May-1998  
 C:Accession: PQ0507; S15488  
 R:Qu, X.; Chen, W.; Maguire, T.; Austin, F.  
 J. Gen. Virol. 74, 89-97, 1993  
 A:Title: Immunoreactivity and protective effects in mice of a recombinant dengue 2 Tonga  
 A:Reference number: PQ0507  
 A:Accession: PQ0507  
 A:Molecule type: mRNA  
 A:Residues: 1-614 <OUX>  
 A:Experimental source: strain Tonga 251  
 R:Qu, X.; Chen, W.; Maguire, T.  
 Submitted to the EMBL Data Library, March 1991  
 A:Description: Nucleotide sequence of the non-structural genes NS1 and NS2a from Dengue-  
 A:Reference number: S15488  
 A:Accession: S15488  
 A:Status: preliminary  
 A:Molecule type: genomic RNA  
 A:Residues: 42-611 <OU2>  
 A:Cross-references: EMBL:X57469; NID:g59311; PID:g930021  
 A:Experimental source: strain Tonga 1974  
 C:Genetics:  
 A:Gene: NS1; NS2A; NS2B  
 C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology  
 C:Keywords: glycoprotein; nonstructural protein NS1 #status predicted <NS1>  
 F:42-393/Product: nonstructural protein NS1 #status predicted <NS1>  
 F:394-611/Product: nonstructural protein NS2a #status predicted <N2A>  
 F:612-614/Product: nonstructural protein NS2b (fragment) #status predicted <N2B>  
 F:171,248,400,440/Blinding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 59.2%; Score 45; DB 2; Length 614;  
 Best Local Similarity 60.0%; Pred. No. 4.4;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 |||:|:|:|  
 Db 48 SWKNKELKCG 57

RESULT 10  
 S33720  
 NS1 protein - dengue virus type 2 (fragment)  
 C:Species: dengue virus type 2  
 C:Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 29-May-1998  
 C:Accession: S33720  
 R:Leblouis, H.; Young, P.R.  
 Nucleic Acids Res. 21, 1668, 1993  
 A:Title: Sequence of the dengue virus type 2 (strain PR-159) NS1 gene and comparison with  
 A:Reference number: S33720

A:Accession: S33720  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-380 <LEB>  
 A:Cross-references: EMBL:Z17213; NID:g288847; PID:g288848  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992  
 C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology

Query Match 59.2%; Score 45; DB 2; Length 380;  
 Best Local Similarity 60.0%; Pred. No. 2.8;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 |||:|:|:|  
 Db 35 SWKNKELKCG 44

RESULT 11  
 B43763  
 nonstructural protein NS1 - dengue virus type 2 (strain D80-030)  
 C:Species: dengue virus type 2  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 29-May-1998  
 C:Accession: B43763  
 R:Blok, J.; Gibbs, A.J.; McWilliam, S.M.; Vitarana, U.T.  
 Arch. Virol. 118, 209-223, 1991  
 A:Title: NS1 gene sequences from eight dengue-2 viruses and their evolutionary relati  
 A:Reference number: A43763  
 A:Accession: B43763  
 A:Molecule type: genomic RNA  
 A:Residues: 1-352 <BLO>  
 A:Cross-references: GB:M58487; NID:g323453; PID:g323454  
 C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology  
 C:Keywords: nonstructural protein

Query Match 59.2%; Score 45; DB 2; Length 352;  
 Best Local Similarity 60.0%; Pred. No. 2.6;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 |||:|:|:|  
 Db 7 SWKNKELKCG 16

RESULT 12  
 C43763  
 nonstructural protein NS1 - dengue virus type 2 (strain PUO-280)  
 C:Species: dengue virus type 2  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 29-May-1998  
 C:Accession: C43763  
 R:Blok, J.; Gibbs, A.J.; McWilliam, S.M.; Vitarana, U.T.  
 Arch. Virol. 118, 209-223, 1991  
 A:Title: NS1 gene sequences from eight dengue-2 viruses and their evolutionary relati  
 A:Reference number: A43763  
 A:Accession: C43763  
 A:Molecule type: genomic RNA  
 A:Residues: 1-352 <BLO>  
 A:Cross-references: GB:M58488; NID:g323455; PID:g323456  
 C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology  
 C:Keywords: nonstructural protein

Query Match 59.2%; Score 45; DB 2; Length 352;  
 Best Local Similarity 60.0%; Pred. No. 2.6;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 |||:|:|:|  
 Db 7 SWKNKELKCG 16

RESULT 13

Query Match	59.2%;	Score 45;	DB 1;	Length 1127;	
Best Local Similarity	60.0%;	Pred. No. 7.7;	1;	Indels	0;
Matches	6;	Conservative	3;	Mismatches	0;
QY	3	SWKSSFIRCQ 12			
DB	782	SWKNKELKCG 791			
: : :					
RESULT	7				
JS0219					
polyprotein - dengue virus type 2 (strain New Guinea-C)					
N;Contains: capsid protein; envelope protein; membrane glycoprotein; nonstructural pr					
n NS4A; nonstructural protein NS4B; nonstructural protein NS5					
C:Species: dengue virus type 2					
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998					
C:Accession: JS0219; A28646					
R;Irie, K.; Mohan, P.M.; Sasaguri, Y.; Putnak, R.; Padmanabhan, R.					
Gene 75, 197-211, 1989					
A:title: Sequence analysis of cloned dengue virus type 2 genome (New Guinea-C strain)					
A:Reference number: JS0219; MUID:89232751					
A:Accession: JS0219					
A:Molecule type: genomic RNA					
A:Residues: 1-3391 <IR>					
R;Putnak, J.R.; Charles, P.C.; Padmanabhan, R.; Irie, K.; Hoke, C.H.; Burke, D.S.					
Virology 163, 93-103, 1988					
A:title: Functional and antigenic domains of the dengue-2 virus nonstructural glycopr					
A:Reference number: A28646; MUID:88160069					
A:Accession: A28646					
A:Molecule type: genomic RNA					
A:Residues: 749-1227 <PUT>					
C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology					
C:Keywords: App; envelope protein; glycoprotein; nonstructural protein; P-loop; polyp					
F:2-114/Product: capsid protein #status predicted <CAP>					
F:115-280/Product: membrane glycoprotein #status predicted <MEM>					
F:115-205/Domain: nonterminal signal sequence #status predicted <SIG>					
F:206-280/Product: membrane glycoprotein #status predicted <MEB>					
F:281-775/Product: envelope protein #status experimental <ENV>					
F:755-775/Domain: transmembrane #status predicted <TM1>					
F:776-1127/Product: nonstructural protein NS1 #status experimental <NS1>					
F:1128-1345/Product: nonstructural protein NS2A #status predicted <NS2>					
F:1135-1146/Domain: transmembrane #status predicted <TM2>					
F:1158-1173/Domain: transmembrane #status predicted <TM3>					
F:1346-1475/Product: nonstructural protein NS2B #status predicted <NSB>					
F:1476-2093/Product: nonstructural protein NS3 #status experimental <NS3>					
F:1668-1675/Domain: DEAD/H box helicase homology <DEAD>					
F:1668-1675/Region: nucleotide-binding motif A (P-loop)					
F:1755-1760/Region: nucleotide-binding motif B					
F:1759-1762/Region: DEAD motif					
F:1794-2379/Product: nonstructural protein NS4A #status predicted <NS4>					
F:2094-2379/Product: nonstructural protein NS4B #status predicted <NS4B>					
F:2380-2491/Product: nonstructural protein NS5 #status experimental <NS5>					
F:2492-3391/Product: nonstructural protein NS5 #status experimental <NS5>					
F:183, 905, 982, 2305, 2457, 2704/Binding site: carbohydrate (Asn) (covalent) #status exper					
F:347, 433, 1134, 1174, 2301, 2485, 2665, 2714/Binding site: carbohydrate (Asn) (covalent) #					
QY	3	SWKSSFIRCQ 12			
DB	782	SWKNKELKCG 791			
: : :					
RESULT	8				
S34368					
nonstructural protein - dengue virus type 2 (fragment)					
C:Species: dengue virus type 2					
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-May-1998					
C:Accession: S34368					
R;Shiu, S.Y.W.; Ip, K.W.Y.; Chan, K.M.					
submitted to the EMBL Data Library, November 1992					
Query Match	59.2%;	Score 45;	DB 2;	Length 3391;	
Best Local Similarity	60.0%;	Pred. No. 21;			
Matches	6;	Conservative	3;	Mismatches	1;
				Indels	0;
				Gaps	0;

Query Match 59.2%; Score 45; DB 1; Length 3388;  
 Best Local Similarity 60.0%; Pred. No. 21;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SWKSSEIRCG 12  
 |||: |||:  
 Db 782 SWKNKELKCG 791

RESULT 3  
 GNMVJA  
 genome polypeptide - dengue virus type 2 (strain Jamaica)  
 N:Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
 C:Species: dengue virus type 2  
 C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 15-May-1998  
 C:Accession: A94346; A94378; A25613; A29199  
 R:Deubel, V.; Kinney, R.M.; Trent, D.W.  
 Virolgy 155, 365-377, 1986  
 A:Title: Nucleotide sequence and deduced amino acid sequence of the structural proteins  
 A:Reference number: A94346; MUID:87071658  
 A:Accession: A94346  
 A:Molecule type: genomic RNA  
 A:Residues: 1-791 <DE1>  
 A:CROSS-references: GB:M15975  
 R:Deubel, V.; Kinney, R.M.; Trent, D.W.  
 Virolgy 165, 234-244, 1988  
 A:Title: Nucleotide sequence and deduced amino acid sequence of the nonstructural proteins  
 A:Reference number: A94378; MUID:88265864  
 A:Accession: A94378  
 A:Molecule type: genomic RNA  
 A:Residues: 792-3391 <DE2>  
 A:CROSS-references: GB:M20558  
 C:Superfamily: yellow fever virus genome polypeptide; DEAD/H box helicase homology  
 C:Keywords: ATP; capsid protein C #status predicted <CPC>  
 F:114/Product: capsid protein C #status predicted <CPC>  
 F:43-59/Domain: transmembrane #status predicted <TM1>  
 F:101-117/Domain: transmembrane #status predicted <TM2>  
 F:115-280/Product: membrane-associated protein M precursor #status predicted <MP>  
 F:115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
 F:206-280/Product: membrane-associated protein M #status predicted <MP>  
 F:268-284/Domain: transmembrane #status predicted <TM3>  
 F:281-775/Product: envelope protein E #status predicted <EPE>  
 F:727-743/Domain: transmembrane #status predicted <TM4>  
 F:757-773/Domain: transmembrane #status predicted <TM5>  
 F:776-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
 F:1128-1345/Product: nonstructural protein NS2a #status predicted <NS2a>  
 F:1346-1474/Product: nonstructural protein NS2b #status predicted <NS2b>  
 F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
 F:1668-1938/Domain: DEAD/H box helicase homology <DEAD>  
 F:1668-1675/Region: nucleotide-binding motif A (P-loop)  
 F:1755-1762/Region: nucleotide-binding motif B  
 F:1759-1762/Region: DEAD motif  
 F:2094-2243/Product: nonstructural protein NS4a #status predicted <NS4a>  
 F:2244-2491/Product: nonstructural protein NS4b #status predicted <NS4b>  
 F:2492-3391/Product: nonstructural protein NS5 #status predicted <NS5>  
 F:183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 59.2%; Score 45; DB 1; Length 3391;  
 Best Local Similarity 60.0%; Pred. No. 21;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SWKSSEIRCG 12  
 |||: |||:  
 Db 782 SWKNKELKCG 791

RESULT 4  
 GNMV16  
 genome polypeptide - dengue virus type 2 (strain 16681)  
 N:Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru

tural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
 C:Species: dengue virus type 2  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 15-May-1998  
 C:Accession: A42451; A43496; A43763  
 R:Block, J.; McWilliam, S.M.; Butler, H.C.; Gibbs, A.J.; Weiller, G.; Herring, B.L.; H  
 Virolgy 187, 573-590, 1992  
 A:Title: Comparison of a dengue-2 virus and its candidate vaccine derivative: sequenc  
 A:Reference number: A42451; MUID:92188532  
 A:Accession: A42451  
 A:Molecule type: genomic RNA  
 A:Residues: 1-3391 <BLO>  
 A:CROSS-references: GB:M84727; GB:M85259; NID:9323472; PID:9323473  
 C:Superfamily: yellow fever virus genome polypeptide; DEAD/H box helicase homology  
 C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protei  
 F:114/Product: capsid protein C #status predicted <CPC>  
 F:115-280/Product: membrane-associated protein M precursor #status predicted <MP>  
 F:115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
 F:206-280/Product: membrane-associated protein M #status predicted <MP>  
 F:268-284/Domain: transmembrane #status predicted <TM1>  
 F:281-775/Product: envelope protein E #status predicted <EPE>  
 F:727-743/Domain: transmembrane #status predicted <TM2>  
 F:757-773/Domain: transmembrane #status predicted <TM3>  
 F:776-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
 F:1128-1345/Product: nonstructural protein NS2a #status predicted <NS2a>  
 F:1346-1474/Product: nonstructural protein NS2b #status predicted <NS2b>  
 F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
 F:1668-1938/Domain: DEAD/H box helicase homology <DEAD>  
 F:1668-1675/Region: nucleotide-binding motif A (P-loop)  
 F:1755-1762/Region: nucleotide-binding motif B  
 F:1759-1762/Region: DEAD motif  
 F:2094-2243/Product: nonstructural protein NS4a #status predicted <NS4a>  
 F:2244-2491/Product: nonstructural protein NS4b #status predicted <NS4b>  
 F:2492-3391/Product: nonstructural protein NS5 #status predicted <NS5>  
 F:183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 59.2%; Score 45; DB 1; Length 3391;  
 Best Local Similarity 60.0%; Pred. No. 21;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SWKSSEIRCG 12  
 |||: |||:  
 Db 782 SWKNKELKCG 791

RESULT 5  
 GNMV26  
 genome polypeptide - dengue virus type 2 (strain 16681-PDK53)  
 N:Contains: capsid protein C; envelope protein E; membrane-associated protein M; nons  
 tural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
 C:Species: dengue virus type 2  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 15-May-1998  
 C:Accession: B42451  
 R:Block, J.; McWilliam, S.M.; Butler, H.C.; Gibbs, A.J.; Weiller, G.; Herring, B.L.; H  
 Virolgy 187, 573-590, 1992  
 A:Title: Comparison of a dengue-2 virus and its candidate vaccine derivative: sequenc  
 A:Reference number: A42451; MUID:92188532  
 A:Accession: B42451  
 A:Molecule type: genomic RNA  
 A:Residues: 1-3391 <BLO>  
 A:CROSS-references: GB:M85259  
 C:Superfamily: yellow fever virus genome polypeptide; DEAD/H box helicase homology  
 C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protei  
 F:114/Product: capsid protein C #status predicted <CPC>  
 F:50-66/Domain: transmembrane #status predicted <TM1>  
 F:102-118/Domain: transmembrane #status predicted <TM2>  
 F:115-280/Product: membrane-associated protein M precursor #status predicted <MP>  
 F:115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
 F:206-280/Product: membrane-associated protein M #status predicted <MP>  
 F:268-284/Domain: transmembrane #status predicted <TM3>  
 F:281-775/Product: envelope protein E #status predicted <EPE>  
 F:727-743/Domain: transmembrane #status predicted <TM4>  
 F:757-773/Domain: transmembrane #status predicted <TM5>

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:43 : Search time 105.14 Seconds  
(without alignments)  
4.954 Million cell updates/sec

Title: US-09-124-280A-29  
Perfect score: 76  
Sequence: 1 CQSWKSSEIRCGK 13

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: Pirl:\*  
2: Pirl2:\*  
3: Pirl3:\*  
4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	136	2 A39497	endotoxin-binding
2	45	59.2	3388	1 GNVWDP	genome polyprotein
3	45	59.2	3391	1 GNVVJA	genome polyprotein
4	45	59.2	3391	1 GNVV16	genome polyprotein
5	45	59.2	3391	1 GNVV26	genome polyprotein
6	45	59.2	1127	1 GNVWD2	genome polyprotein
7	45	59.2	3391	2 JS0219	polyprotein - deng
8	45	59.2	352	2 S34368	nonstructural prot
9	45	59.2	614	2 PQ0507	polyprotein - deng
10	45	59.2	380	2 S33720	NS1 protein - deng
11	45	59.2	352	2 B43763	nonstructural prot
12	45	59.2	352	2 C43763	nonstructural prot
13	45	59.2	352	2 D43763	nonstructural prot
14	45	59.2	352	2 E43763	nonstructural prot
15	45	59.2	352	2 F43763	nonstructural prot
16	45	59.2	352	2 G43763	nonstructural prot
17	45	59.2	352	2 H43763	nonstructural prot
18	42	55.3	659	2 S74435	arginine decarboxy
19	42	55.3	352	2 S09242	nonstructural prot
20	41	53.9	3390	1 GNVWD3	genome polyprotein
21	40	52.6	164	1 NNMU3	2S albumin 3 precu
22	40	52.6	3396	1 AA2551	genome polyprotein
23	40	52.6	1226	1 GNVWVP	18K lipopolysaccha
24	40	52.6	171	2 JO1171	genome polyprotein
25	40	52.6	473	2 S37467	genome polyprotein
26	40	52.6	473	2 S37470	genome polyprotein
27	39	51.3	122	1 PSABA	phospholipase A2 (
28	39	51.3	106	1 L4HUM1	ig lambda chain v-
29	39	51.3	3386	1 GNVWDF	genome polyprotein
30	39	51.3	864	1 VCLJGA	env polyprotein -
31	39	51.3	364	2 S31304	protein-tyrosine-p
32	39	51.3	337	2 B71052	hypothetical prote
33	39	51.3	31	2 C26871	hypothetical chl p
34	39	51.3	652	2 T02001	hypothetical prote
35	39	51.3	414	2 S42159	lipic acid syntha
36	38.5	50.7	695	2 S05008	serine proteinase,
37	38.5	50.7	407	2 T00693	hypothetical prote
38	38	50.0	329	1 VMDT21	VSG expression sit
39	38	50.0	330	2 B42941	site-specific DNA-

ALIGNMENTS

RESULT 1

A39497  
endotoxin-binding protein/proteinase inhibitor precursor - Atlantic horseshoe crab  
C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
C:Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 09-Sep-1997  
C:Accession: A39497  
R:Minetti, C.A.S.A.; Lin, Y.; Cislowski, T.; Liu, T.Y.  
J. Biol. Chem. 266, 20773-20780, 1991  
A:Title: Purification and characterization of an endotoxin-binding protein with prote  
A:Reference number: A39497; MUID:92041932  
A:Accession: A39497  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-136 <MIN>  
A:CROSS-references: GB:M65017; NID:gl56637; PID:gl56638

Query Match 100.0% Score 76; DB 2; Length 136;  
Best Local Similarity 100.0%; Pred. No. 7.3e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQSWKSSEIRCGK 13  
|||||  
DB 36 CQSWKSSEIRCGK 48

RESULT 2

GNVWDP  
genome polyprotein - dengue virus type 2 (strain PRI59/S1)  
N:Contains: capsid protein; envelope protein; membrane protein; nonstructural protein  
a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: dengue virus type 2  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 15-May-1998  
C:Accession: A29972  
R:Hahn, Y.S.; Geller, R.; Hunkapiller, T.; Dalrymple, J.M.; Strauss, J.H.; Strauss, E  
Virology 162, 167-180, 1988  
A:Title: Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins w  
A:Reference number: A29972; MUID:88101365  
A:Accession: A29972  
A:Molecule type: genomic RNA  
A:Residues: 1-3388 <HAH>  
A:CROSS-references: GB:M19197; NID:g323654; PID:g323655  
C:Superfamily: yellow fever virus genome polyprotein; DEAD/H box helicase homology  
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; membrane protein; no  
F:2-114/Product: capsid protein #status predicted <CAP>  
F:115-280/Product: membrane protein precursor #status predicted <MPP>  
F:115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
F:206-280/Product: membrane protein #status predicted <MMP>  
F:281-775/Product: envelope protein #status predicted <ENP>  
F:776-1186/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1189-1345/Product: nonstructural protein NS2a #status predicted <N2A>  
F:1346-1475/Product: nonstructural protein NS2b #status predicted <N2B>  
F:1476-2090/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1668-1938/Domain: DEAD/H box helicase homology <DEAD>  
F:1668-1675/Region: nucleotide-binding motif A (P-loop)  
F:1755-1760/Region: nucleotide-binding motif B  
F:1759-1762/Region: DEAD motif  
F:2091-2376/Product: nonstructural protein NS4a #status predicted <N4A>  
F:2377-2488/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2489-3388/Product: nonstructural protein NS5 #status predicted <NS5>  
F:183,347,433,905,982,1134,1174,1329,1369,2298,2302,2384,2454,2482,2641,2662,2701,271

**THIS PAGE BLANK (USPTO)**

-Wed Sep 8 09:43:50 1999

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/216,594  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 1795/39318  
TELEPHONE: 212-977-9550  
TELEFAX: 212-664-0525  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 390 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
IMMEDIATE SOURCE:  
CLONE: 5-HT1DB  
US-08-216-594-6

Query Match 50.0%; Score 38; DB 1; Length 390;  
Best Local Similarity 54.5%; Pred. No. 62;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CQSWKSSEIRC 11  
DB 122 CDFWLSDDITC 132

Search completed: September 7, 1999, 23:07:48  
Job time: 1746 sec

Query Match 50.0%; Score 38; DB 1; Length 376;  
Best Local Similarity 54.5%; Pred. No. 60;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 COSWKSSEIRC 11  
| | | | | | | | | |  
Db 111 CDWLSSDITC 121

RESULT 13  
US-08-117-006-6  
; Sequence 6, Application US/08117006  
; Patent No. 5639652  
; GENERAL INFORMATION:  
; APPLICANT: Weinshank, Richard L.  
; APPLICANT: Branchek, Theresa  
; APPLICANT: Hartig, Paul R.  
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect, Version 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/117,006  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 1795/39318  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-664-0525  
; TELEFAX: 212-664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 390 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; IMMEDIATE SOURCE:  
; CLONE: 5-HT1DB  
US-08-117-006-6

Query Match 50.0%; Score 38; DB 1; Length 390;  
Best Local Similarity 54.5%; Pred. No. 62;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 COSWKSSEIRC 11  
| | | | | | | | | |  
Db 122 CDFWLSSDITC 132

RESULT 14  
US-08-216-594-5  
; Sequence 5, Application US/08216594  
; Patent No. 5652113  
; GENERAL INFORMATION:

; APPLICANT: Weinshank, Richard L.  
; APPLICANT: Branchek, Theresa  
; APPLICANT: Hartig, Paul R.  
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/216,594  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 1795/39318  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-977-9550  
; TELEFAX: 212-664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 376 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; IMMEDIATE SOURCE:  
; CLONE: 5-HT1DA  
US-08-216-594-5

Query Match 50.0%; Score 38; DB 1; Length 376;  
Best Local Similarity 54.5%; Pred. No. 60;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 COSWKSSEIRC 11  
| | | | | | | | | |  
Db 111 CDWLSSDITC 121

RESULT 15  
US-08-216-594-6  
; Sequence 6, Application US/08216594  
; Patent No. 5652113  
; GENERAL INFORMATION:  
; APPLICANT: Weinshank, Richard L.  
; APPLICANT: Branchek, Theresa  
; APPLICANT: Hartig, Paul R.  
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:

STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118,270  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-29

Query Match 50.08; Score 38; DB 1; Length 327;  
Best Local Similarity 54.58; Pred. No. 52;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 COSWKSSEIRC 11  
| | | | |  
Db 73 CDFWLSSDITC 83

RESULT 11  
US-08-118-270-226  
Sequence 226, Application US/08118270  
Patent No. 5508384  
GENERAL INFORMATION:  
APPLICANT: Murphy, Randall B.  
APPLICANT: Schuster, David I.  
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118,270  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 226:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 41 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-226  
Query Match 50.08; Score 38; DB 1; Length 41;  
Best Local Similarity 54.58; Pred. No. 6.5;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 COSWKSSEIRC 11  
| | | | |  
Db 18 CDFWLSSDITC 28

RESULT 12  
US-08-117-006-5  
Sequence 5, Application US/08117006  
Patent No. 5639652  
GENERAL INFORMATION:  
APPLICANT: Weinsbank, Richard L.  
APPLICANT: Branchek, Theresa  
APPLICANT: Hartig, Paul R.  
TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR AND  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect, Version 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/117,006  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 1795/39318  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-977-9550  
TELEFAX: 212-664-0525  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
IMMEDIATE SOURCE:  
CLONE: 5-HT1DA  
US-08-117-006-5



Best Local Similarity 54.5%; Pred. No. 62;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 COSWKSSEIRC 11  
Db 122 CDFWLSSDITC 132

## RESULT 8

US-08-370-542-5  
; Sequence 5, Application US/08370542  
; Patent No. 5476782  
; GENERAL INFORMATION:  
; APPLICANT: Weinshank, Richard L.  
; APPLICANT: Branchek, Theresa  
; APPLICANT: Hartig, Paul R.  
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1E RECEPTOR AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370.542  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/194.113  
; FILING DATE:  
; APPLICATION NUMBER: US/07/803.626  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 1795/39317  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-977-9550  
; TELEFAX: 212-664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 375 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
US-08-370-542-5

Query Match 50.0%; Score 38; DB 1; Length 375;  
Best Local Similarity 54.5%; Pred. No. 60;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 COSWKSSEIRC 11  
Db 111 CDFWLSSDITC 121

## RESULT 9

US-08-370-542-6  
; Sequence 6, Application US/08370542

; Patent No. 5476782  
; GENERAL INFORMATION:  
; APPLICANT: Weinshank, Richard L.  
; APPLICANT: Branchek, Theresa  
; APPLICANT: Hartig, Paul R.  
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1E RECEPTOR AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370.542  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/194.113  
; FILING DATE:  
; APPLICATION NUMBER: US/07/803.626  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 1795/39317  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-977-9550  
; TELEFAX: 212-664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 398 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
US-08-370-542-6

Query Match 50.0%; Score 38; DB 1; Length 398;  
Best Local Similarity 54.5%; Pred. No. 64;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 COSWKSSEIRC 11  
Db 122 CDFWLSSDITC 132

## RESULT 10

US-08-118-270-29  
; Sequence 29, Application US/08118270  
; Patent No. 5508384  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Randall B.  
; APPLICANT: Schuster, David I.  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington



;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;;  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;;  
;; INFORMATION FOR SEQ ID NO: 29:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 13 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;;  
;; US-08-456-112B-29

Query Match 100.0%; Score 76; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 COSKSSSEIRCGK 13  
|||||

DB 1 COSKSSSEIRCGK 13  
|||||

RESULT 3  
PCT-US96-07627-3  
;; Sequence 3, Application PC/TUS9607627  
;; GENERAL INFORMATION:  
;; APPLICANT:  
;; TITLE OF INVENTION: SUBUNIT VACCINE AGAINST FLAVIVIRUS INFECTION  
;; NUMBER OF SEQUENCES: 50  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US96/07627  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1127 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
;; PCT-US96-07627-3

Query Match 59.2%; Score 45; DB 3; Length 1127;  
Best Local Similarity 60.0%; Pred. No. 15;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSSEIRCG 12  
|||

DB 782 SWKNKELCKG 791  
|||

RESULT 4  
US-09-057-762-13  
;; Sequence 13, Application US/09057762  
;; Patent No. 5879909  
;; GENERAL INFORMATION:  
;; APPLICANT: PERL, ANDRAS  
;; TITLE OF INVENTION: HUMAN TRANSALDOLASE: AN AUTOANTIGEN WITH  
;; TITLE OF INVENTION: A FUNCTION IN METABOLISM  
;; NUMBER OF SEQUENCES: 24  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: MORRISON & FOERSTER  
;; STREET: 2000 Pennsylvania Avenue N.W.  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20006-1812  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/057,762  
;; FILING DATE:  
;; CLASSIFICATION: 530  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/326,119  
;; FILING DATE: 19-OCT-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: LIVNAT, SHMUEL  
;; REGISTRATION NUMBER: 33,949  
;; REFERENCE/DOCKET NUMBER: 280932000100  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 887-1500  
;; TELEFAX: (202) 822-0168  
;;  
;; INFORMATION FOR SEQ ID NO: 13:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 138 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;;  
;; US-09-057-762-13

Query Match 53.9%; Score 41; DB 2; Length 138;  
Best Local Similarity 50.0%; Pred. No. 7.6;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SWKSSSEIRCG 12  
|||

DB 16 NWKGKELCKG 25  
|||

RESULT 5  
US-08-313-681A-4  
;; Sequence 4, Application US/08313681A  
;; Patent No. 5618675  
;; GENERAL INFORMATION:  
;; APPLICANT: Larrick, James W.  
;; APPLICANT: Wright, Susan C.  
;; APPLICANT: Hirata, Mishimasa  
;; TITLE OF INVENTION: Human Cationic Proteins Having  
;; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity  
;; NUMBER OF SEQUENCES: 30  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend Kourie and Crew  
;; STREET: One Market Plaza, Steuart tower, Suite 2000  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: USA

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:47 ; Search time 82.37 Seconds  
(without alignments)  
1.557 Million cell updates/sec

Title: US-09-124-280A-29

Perfect score: 76

Sequence: 1 COSWKSSEIRCGK 13

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	13	1	US-08-097-830E-29
2	76	100.0	13	2	US-08-456-112B-29
3	45	59.2	1127	3	PCT-US96-07627-3
4	41	53.9	138	1	US-09-057-762-13
5	40	52.6	171	1	US-08-313-681A-4
6	38	50.0	376	1	US-07-817-920-5
7	38	50.0	390	1	US-07-817-920-6
8	38	50.0	375	1	US-08-370-542-5
9	38	50.0	398	1	US-08-370-542-6
10	38	50.0	327	1	US-08-118-270-29
11	38	50.0	41	1	US-08-118-270-226
12	38	50.0	376	1	US-08-117-006-5
13	38	50.0	390	1	US-08-117-006-6
14	38	50.0	376	1	US-08-216-594-5
15	38	50.0	390	1	US-08-216-594-6
16	38	50.0	375	2	US-08-542-358-5
17	38	50.0	398	2	US-08-542-358-6
18	38	50.0	622	2	US-08-459-146-2
19	38	50.0	622	2	US-08-459-065-2
20	38	50.0	376	3	PCT-US93-00149-5
21	38	50.0	390	3	PCT-US93-00149-6
22	38	50.0	327	3	PCT-US93-08528-29
23	38	50.0	41	3	PCT-US93-08528-226
24	37	48.7	110	2	US-08-482-882-96
25	37	48.7	110	2	US-08-483-389-96
26	37	48.7	110	2	US-08-487-113D-96
27	37	48.7	110	2	US-08-473-503-96
28	37	48.7	110	2	US-08-483-932-96
29	36	47.4	127	2	US-08-482-882-80
30	36	47.4	127	2	US-08-483-389-80
31	36	47.4	127	2	US-08-487-113D-80
32	36	47.4	64	2	US-08-765-179B-19
33	36	47.4	127	2	US-08-473-503-80
34	36	47.4	127	2	US-08-483-932-80
35	35	46.1	201	1	US-08-155-171B-19
36	35	46.1	558	1	US-08-333-358-12
37	35	46.1	107	1	US-08-459-310-14
38	35	46.1	107	1	US-08-459-310-17
39	35	46.1	558	1	US-08-463-694-12

40 35 46.1 558 1 US-08-694-501-12 Sequence 12, Appl  
41 35 46.1 61 2 US-08-682-485A-26 Sequence 26, Appl  
42 34 44.7 483 1 US-08-194-338-7 Sequence 7, Appl  
43 34 44.7 33 1 US-08-428-248-2 Sequence 2, Appl  
44 34 44.7 33 1 US-08-428-248-7 Sequence 7, Appl  
45 34 44.7 20 3 PCT-US95-09307-19 Sequence 19, Appl

ALIGNMENTS

RESULT 1  
US-08-097-830E-29  
; Sequence 29, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-29

Query Match 100.0% Score 76; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 COSWKSSEIRCGK 13  
Db 1 COSWKSSEIRCGK 13

RESULT 2  
US-08-456-112B-29  
; Sequence 29, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45

**THIS PAGE BLANK (USPTO)**

AC R15137;  
 DT 17-FEB-1992 (first entry)  
 DE Human serotonin 1D receptor encoded by gene 8-30-84.  
 KW 5-HT (1D); Parkinson's Disease; migraine; anxiety; eating disorder;  
 KW G-protein; 5-hydroxytryptamine.  
 OS Homo sapiens.  
 FH Key  
 FT domain  
 FT 1. .38  
 FT /label= extracellular  
 FT 39. .62  
 FT /label= transmembrane-1  
 FT 63. .75  
 FT /label= cytoplasmic  
 FT 76. .99  
 FT /label= transmembrane-2  
 FT 100. .110  
 FT /label= extracellular  
 FT 111. .134  
 FT /label= transmembrane-3  
 FT 135. .153  
 FT /label= cytoplasmic  
 FT 154. .177  
 FT /label= transmembrane-4  
 FT 178. .194  
 FT /label= extracellular  
 FT 195. .218  
 FT /label= transmembrane-5  
 FT 219. .301  
 FT /label= cytoplasmic\_loop  
 FT 302. .325  
 FT /label= transmembrane-6  
 FT 326. .334  
 FT /label= extracellular  
 FT 335. .358  
 FT /label= transmembrane-7  
 FT 359. .376  
 FT /label= cytoplasmic\_tail  
 FT W09117174-A.  
 PN 14-NOV-1991.  
 PD 08-MAY-1991; U03200.  
 PF 08-MAY-1991; US-520716.  
 PR (NEUR-) NEUROGENETIC CORP.  
 PA Weinshank RL, Branchek T, Hartig PR;  
 PI WPI; 91-353715/48.  
 DR N-PSDB: Q14835.  
 DT Nucleic acid encoding 5HT-1D receptors and their antibodies -  
 PT used to treat and diagnose conditions caused by abnormal 5HT-1D  
 PT receptor expression e.g. dementia  
 PS Claim 6; Fig 3; 90pp; English.  
 CC There are two subtypes of 5HT(1D) receptors encoded by gene 8-30-84  
 CC and gene 11, respectively. A full-length clone corresponding to  
 CC gene 8-30-84 was isolated from a human hippocampus library. The  
 CC amino acid sequence deduced from the nucleotide sequence predicts a  
 CC mol.wt. 41,783 protein. Comparison of this protein sequence with  
 CC previously characterised neurotransmitter receptors suggests that  
 CC it is a new member of the G protein-coupled receptor family.  
 CC See Q14836 for gene 11.  
 SQ Sequence 376 AA;

Query Match 50.0%; Score 38; DB 1; Length 376;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Qy 1 COSWKSSEIRC 11  
 Db 111 CDWLSSDITC 121

RESULT 15  
 R15138  
 ID R15138 standard; Protein; 398 AA.  
 AC R15138;

DT 17-FEB-1992 (first entry)  
 DE Human serotonin 1D receptor encoded by gene 11.  
 KW 5-HT (1D); Parkinson's Disease; migraine; anxiety; eating disorder;  
 KW G-protein; 5-hydroxytryptamine.  
 OS Homo sapiens.  
 FH Key  
 FT domain  
 FT 1. .49  
 FT /label= extracellular  
 FT 50. .73  
 FT /label= transmembrane-1  
 FT 74. .86  
 FT /label= cytoplasmic  
 FT 87. .110  
 FT /label= transmembrane-2  
 FT 111. .121  
 FT /label= extracellular  
 FT 122. .145  
 FT /label= transmembrane-3  
 FT 146. .164  
 FT /label= cytoplasmic  
 FT 165. .188  
 FT /label= transmembrane-4  
 FT 189. .205  
 FT /label= extracellular  
 FT 206. .229  
 FT /label= transmembrane-5  
 FT 230. .315  
 FT /label= cytoplasmic\_loop  
 FT 316. .339  
 FT /label= transmembrane-6  
 FT 340. .348  
 FT /label= extracellular  
 FT 349. .372  
 FT /label= transmembrane-7  
 FT 373. .398  
 FT /label= cytoplasmic\_tail  
 FT W09117174-A.  
 PN 14-NOV-1991.  
 PD 08-MAY-1991; U03200.  
 PF 08-MAY-1991; US-520716.  
 PR (NEUR-) NEUROGENETIC CORP.  
 PA Weinshank RL, Branchek T, Hartig PR;  
 PI WPI; 91-353715/48.  
 DR N-PSDB: Q14836.  
 DT Nucleic acid encoding 5HT-1D receptors and their antibodies -  
 PT used to treat and diagnose conditions caused by abnormal 5HT-1D  
 PT receptor expression e.g. dementia  
 PS Claim 7; Fig 4; 90pp; English.  
 CC There are two subtypes of 5HT(1D) receptors encoded by gene 8-30-84  
 CC and gene 11, respectively. A full-length clone corresponding to  
 CC gene 11 was isolated from a human placental library. The  
 CC amino acid sequence deduced from the nucleotide sequence predicts a  
 CC mol.wt. 44,333 protein. Comparison of this protein sequence with  
 CC previously characterised neurotransmitter receptors suggests that  
 CC it is a new member of the G protein-coupled receptor family.  
 CC See Q14836 for gene 11.  
 SQ Sequence 398 AA;

Query Match 50.0%; Score 38; DB 1; Length 398;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Qy 1 COSWKSSEIRC 11  
 Db 122 CDFWLSSDITC 132

Search completed: September 7, 1999, 22:50:01  
 Job time: 7882 sec

FT PN WO9322440-A. /label= NS5  
PD 11-NOV-1993.  
PF 28-APR-1993; CA0182.  
PR 29-APR-1992; GB-009243.  
PA (UYSI-) UNIV SINGAPORE NAT.  
PI Chan Y, Fu J, Tan B, Yap E, Tan Y;  
DR WPI: 93-368799/46.  
DR N-PSDB: 051476.  
PT New Dengue virus type 1 strain - used to obtain prods. for  
PT detection, diagnosis, vaccines and treatment involving virus  
PS Claim 9: Page 20-34; 53pp; English.  
CC DEN1 virus, strain S275/90 was isolated from the serum of a dengue  
CC haemorrhagic fever (DHF) patient. RNA was isolated from the virus  
CC and used to prepare cDNA encoding DEN1 polyproteins. Dengue virus  
CC Type 1 prods. can be used for detection, diagnosis, vaccines  
CC (inactivated form) or treatment of DEN1 infections. The sequences  
CC given in 051477-86 are oligonucleotides used to prepare cDNA  
CC fragments corresp. to dengue virus proteins, by PCR.  
SQ Sequence 3396 AA;

Query Match 52.6%; Score 40; DB 1; Length 3396;  
Best Local Similarity 50.0%; Pred. No. 5.4e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 3 SWKSSEIRCG 12  
:|:|:|:|  
Db 781 NWKGRELKCG 790

RESULT 12  
R88719  
ID R88719 standard; protein; 108 AA.  
AC R88719;  
DE Human antibody lambda chain variable region consensus sequence.  
KW Antibody; stability; instability; mutagenesis; amino acid frequency;  
KW canonical sequence approximation; substitution; immunoglobulin;  
KW light chain; variable region; VL-region; diagnosis; immunosay;  
KW cancer treatment; autoimmune disease; immunotoxin; improved yield.  
OS Homo sapiens.  
PN DE4425115-A1.  
PD 18-JAN-1996.  
PF 15-JUL-1994; 425115.  
PR 15-JUL-1994; DE-425115.  
PA (BOEF) BOEHRINGER MANNHEIM GMBH.  
FI Steinbacher S, Stelpe B;  
DR WPI: 96-069594/08.  
PT Functional antibodies modified to increase or decrease stability -  
PT by reference to consensus sequences  
PS Claim 2; Page 31; 43pp; German.  
CC Consensus sequences were established for antibody variable regions  
CC from human and mouse heavy and light (kappa and lambda) chains. In  
CC addition, frequency tables listing the amino acids which can be  
CC found at each position were compiled. In a new method, mutations are  
CC introduced into an antibody variable domain by reference to the  
CC frequency table. When an amino acid is replaced by one which occurs  
CC at a higher frequency at that position, the resulting antibody chain  
CC is more stable than the wild-type; when an amino acid is replaced by  
CC one which occurs less frequently (or not at all), the resulting  
CC antibody is less stable than the wild-type. Stabilised antibodies  
CC are useful as diagnostic reagents, as catalysts and in treatment of  
CC cancer, autoimmune diseases and infections. Destabilised antibodies  
CC have improved (faster) pharmacokinetic properties.  
CC The present sequence is the human lambda light chain variable region  
CC consensus sequence.  
SQ Sequence 108 AA;

Query Match 52.6%; Score 40; DB 1; Length 108;  
Best Local Similarity 58.3%; Pred. No. 16;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 CQSWKSSEIRCG 12  
:|:|:|:|:|  
Db 87 CQSWSSSVVFG 98

RESULT 13  
W06592  
ID W06592 standard; Protein; 798 AA.  
AC W06592;  
DE Amino acid sequence of DEN-1 16007 PDK-13 vaccine virus.  
KW Dengue 2 virus; polyprotein; capsid; prM; M; E; NS1; NS2A; NS2B; NS3;  
KW NS4A; NS4B; NS5; PDK-53; quadravalent vaccine; immunity; serotype;  
KW chimeric DEN-2/1 virus; chimeric DEN-2/3 virus; chimeric DEN-2/4 virus;  
KW dengue fever; fatal dengue haemorrhagic fever; dengue shock syndrome;  
KW DHF; DSS.  
OS Dengue 2 virus.  
OS Synthetic.  
FH Key  
FT Protein  
FT /label= Partial\_capsid\_protein  
FT protein  
FT 84..174  
FT /label= prM  
FT 175..249  
FT /label= M  
FT 250..744  
FT /label= E  
FT 745..798  
FT /label= Partial\_NS1\_protein  
PN W0640933-A1.  
PD 19-DEC-1996.  
PF 05-JUN-1996; U09209.  
PR 07-JUN-1995; US-483292.  
PA (UYMA-) UNIV MAHIDOL AT SALAYA.  
PI Bhamarapavati N, Butrapet S, Chang J, Gubler DJ;  
PI Halstead SB, Kinney R, Trent DW;  
DR WPI: 97-052330/05.  
DR N-PSDB: T49305.  
PT PDK-53, a clone of infectious attenuated Dengue 2 virus strain 16681  
PT - also chimeric DEN-2/1, DEN-2/3 and DEN-1/4 viruses, used as a  
PT quadravalent vaccine for protecting against Dengue virus infection  
PS Claim 31; Page 172-176; 261pp; English.  
CC This sequence represents the amino acid sequence of one of the  
CC candidate vaccine viruses of the invention. This sequence is based  
CC on a fragment of the polyprotein from Dengue 2 virus, strain 16681.  
CC The polyprotein comprises the capsid, prM, M, E, NS1, NS2A, NS2B, NS3,  
CC NS4A, NS4B and NS5 proteins. The quadravalent vaccine of the invention  
CC comprises an attenuated Dengue virus clone, PDK-53, and a chimeric  
CC DEN-2/1 virus, a chimeric DEN-2/3 virus, and/or a chimeric DEN-2/4  
CC virus. The new quadravalent vaccines are used to protect against  
CC infection by all four serotypes of dengue virus, DEN-1, DEN-2, DEN-3  
CC and DEN-4, which can lead to dengue fever or fatal dengue haemorrhagic  
CC fever/dengue shock syndrome (DHF/DSS). Host cells are used to produce  
CC the recombinant protein products of the DNA constructs which are used  
CC in the vaccines.  
SQ Sequence 798 AA;

Query Match 52.6%; Score 40; DB 1; Length 798;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 3 SWKSSEIRCG 12  
:|:|:|:|:|  
Db 751 NWKGRELKCG 760

RESULT 14  
R15137  
ID R15137 standard; Protein; 376 AA.

```

Best Local Similarity 60.0%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
   ||| |::||
Db 782 SWNKELKCG 791

RESULT 9
W06593
ID W06593 standard; Protein; 715 AA.
AC W06593;
DT 15-SEP-1997 (first entry)
DE Amino acid sequence of DEN-3 16562 PGMK-30/FRHL-3 vaccine virus.
KW DEN-3 virus; polypeptide; capsid; prM; M; E; NS1; NS2A; NS2B; NS3;
KW NS4A; NS4B; NS5; PK-53; quadravalent vaccine; immunity; serotype;
KW chimeric DEN-2/1 virus; chimeric DEN-2/3 virus; chimeric DEN-2/4 virus;
KW dengue fever; fatal dengue haemorrhagic fever; dengue shock syndrome;
KW DHF; DSS.
OS Dengue 2 virus.
OS Synthetic.
FH Key
FT Location/Qualifiers
FT 1..36
FT /label= Partial_capsid_protein
FT 37..127
FT /label= prM
FT 128..202
FT /label= M
FT 203..695
FT /label= E
FT 696..715
FT /label= Partial_NS1_protein
FN W06640933-Al.
PD 19-DEC-1996.
PF 06-JUN-1996; U09209.
PR 07-JUN-1995; US-483292.
PA (UYMA-) UNIV MAHIDOL AT SALAYA.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Bhamarapravati N, Rutrapiet S, Chang J, Gubler DJ;
PI Halstead SB, Kinney R, Trent DW;
DR WPI; 97-052330/05.
DR N-PSDB: T49306.
PT PK-53, a clone of infectious attenuated Dengue 2 virus strain 16681
PT - also chimeric DEN-2/1, DEN-2/3 and DEN-1/4 viruses, used as a
PT quadravalent vaccine for protecting against Dengue virus infection
PT Claim 33; Page 176-179; 261pp; English
CC This sequence represents the amino acid sequence of one of the
CC candidate vaccine viruses of the invention. This sequence is based
CC on a fragment of the polypeptide from Dengue 2 virus, strain 16681.
CC The polypeptide comprises the capsid, prM, M, E, NS1, NS2A, NS2B, NS3,
CC NS4A, NS4B and NS5 proteins. The quadravalent vaccine of the invention
CC comprises an attenuated Dengue virus clone, PK-53, and a chimeric
CC DEN-2/1 virus, a chimeric DEN-2/3 virus, and/or a chimeric DEN-2/4
CC virus. The new quadravalent vaccines are used to protect against
CC infection by all four serotypes of dengue virus, DEN-1, DEN-2, DEN-3
CC and DEN-4, which can lead to dengue fever or fatal dengue haemorrhagic
CC fever/dengue shock syndrome (DHF/DSS). Host cells are used to produce
CC the recombinant protein products of the DNA constructs which are used
CC in the vaccines.
SQ Sequence 715 AA;

Query Match 53.9%; Score 41; DB 1; Length 715;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
   ||| |::||
Db 702 NWKELKCG 711

RESULT 10
W72466
ID W72466 standard; peptide; 9 AA.
AC W72466;
DT 23-DEC-1998 (first entry)
DE Dengue virus type-2 glycoprotein NS1 peptide for epitope mapping #11.
KW Dengue virus type-2 glycoprotein NS1; dengue haemorrhagic fever; DHF;
KW dengue shock syndrome; DSS; Aedes aegypti; mosquito; antigen; vaccine;
KW immunisation; immunoreactive; infection.
OS Dengue virus.
OS US5824506-A.
PN 20-OCT-1998.
PD 15-AUG-1994; P. 290268.
PF 15-AUG-1994; US-290268.
PR (GENE-) GENELABS DIAGNOSTICS PTE LTD.
PA Chan L, Guan M;
PI WPI; 98-582557/49.
DR Dengue virus peptide antigens - especially for diagnosis of dengue
PT virus infection
PT Example 1; Column 17; 21pp; English.
PS W72456 to W72570 represent peptide fragments from the dengue virus
CC type-2 glycoprotein NS1, which was used in an example from the present
CC invention for an epitope mapping assay. The invention has developed
CC peptide antigens consisting of fragments of the dengue virus NS1
CC protein. The peptide antigens can be used for the diagnosis of dengue
CC virus infection by detection of antibodies to the virus, especially in
CC an assay comprising attaching the antigen to a solid support, contacting
CC a serum sample with the support, and detecting bound antibodies with a
CC labelled anti-human antibody or used for preparing vaccines against
CC dengue virus infection.
SQ Sequence 9 AA;

Query Match 53.9%; Score 41; DB 1; Length 9;
Best Local Similarity 55.6%; Pred. No. 1.5e+05;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 WKSSEIRCG 12
   ||| |::||
Db 1 WKNKELKCG 9

RESULT 11
R43662
ID R43662 standard; Protein; 3396 AA.
AC R43662;
DT 16-MAY-1994 (first entry)
DE DEN1-S275/90 (BCACC V92042111).
KW Dengue haemorrhagic fever; DHF; dengue fever; DF; dengue shock syndrome;
KW DSS; DEN1 polypeptides.
FH Dengue Fever Virus Type 1 strain S275/90.
FH Key
FT Location/Qualifiers
FT 1..114
FT /label= C
FT 15..114
FT /label= C'
FT 115..205
FT /label= Prem
FT 206..280
FT /label= M
FT 281..774
FT /label= E
FT 775..1128
FT /label= NS1
FT 1129..1344
FT /label= NS2A
FT 1345..1474
FT /label= NS2B
FT 1475..2093
FT /label= NS3
FT 2094..2242
FT /label= NS4A
FT 2243..2492
FT /label= NS4B
FT 2493..3396

```



FT modified\_site 982 /note= "N-linked glycosylation site"

FT protein 1128.1345 /label= NS2A

FT protein 1346.1475 /label= NS2B

FT protein 1476.12093 /label= NS3

FT protein 2094.2242 /label= NS4A

FT protein 2243.2491 /label= NS4B

FT protein 2492.3391 /label= NS5

FT misc\_difference 3038

FT /note= "Encoded by KKA"

PN WO9640933-A1.

PD 19-DEC-1996.

PF 06-JUN-1996; U09209.

PR 07-JUN-1995; US-483292.

PA (USMA) UNIV MAHIDOL AT SALAYA.

PI (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Bhamarapravati N, Butrapet S, Chang J, Gubler DJ;

PI Halstead SB, Kinney R, Trent DW;

DR WPI; 97-052330/05.

DR N-PSDB; T49303.

PT PDK-53, a clone of infectious attenuated Dengue 2 virus strain 16681

PT - also chimeric DEN-2/1, DEN-2/3 and DEN-1/4 viruses, used as a

PT quadravalent vaccine for protecting against Dengue virus infection

PS Claim 23; Page 107-121; 261pp; English.

CC This sequence represents the polyprotein from Dengue 2 virus, strain

CC 16681. The polyprotein comprises the capsid, prM, M, E, NS1, NS2A,

CC NS2B, NS3, NS4A, NS4B and NS5 proteins. A clone of this wildtype

CC viral sequence, PDK-53, may be used in the production of a quadravalent

CC vaccine which provides immunity against all four serotypes of dengue

CC virus. The vaccine also comprises a chimeric DEN-2/1 virus, a chimeric

CC DEN-2/3 virus, and/or a chimeric DEN-2/4 virus. The new quadravalent

CC vaccines are used to protect against infection by all four serotypes of

CC dengue virus, DEN-1, DEN-2, DEN-3 and DEN-4, which can lead to dengue

CC fever or fatal dengue haemorrhagic fever/dengue shock syndrome (DHF/DSS).

CC Host cells are used to produce the recombinant protein products of the

CC DNA constructs which are used in the vaccines.

SQ Sequence 3391 AA;

Query Match 59.2%; Score 45; DB 1; Length 3391;

Best Local Similarity 60.0%; Pred. NO. 91;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12

III: I::II

Db 782 SWKNKELKCG 791

RESULT 7

W72436

ID W72436 standard; protein; 352 AA.

DT 23-DEC-1998 (first entry)

DE Dengue virus type-2 glycoprotein NS1.

KW Dengue virus type-2 glycoprotein NS1; dengue haemorrhagic fever; DHF;

KW dengue shock syndrome; DSS; Aedes aegypti; mosquito; antigen; vaccine;

OS immunisation; immunoreactive; infection.

OS Dengue virus.

PN US5824506-A.

PD 20-OCT-1998.

PF 15-AUG-1994; 290268.

PR 15-AUG-1994; US-290268.

PA (GENE-) GENELABS DIAGNOSTICS PTE LTD.

PI Chan L, Guan M;

DR WPI; 98-582552/49.

PT Dengue virus peptide antigens - especially for diagnosis of dengue

PT virus infection

PS Claim 1; Fig 2; 21pp; English.

CC The present sequence represents the dengue virus type-2 glycoprotein NS1.

CC The present invention has developed peptide antigens consisting of

CC fragments of the dengue virus NS1 protein. The peptide antigens can be

CC used for the diagnosis of dengue virus infection by detection of

CC antibodies to the virus, especially in an assay comprising attaching the

CC antigen to a solid support, contacting a serum sample with the support,

CC and detecting bound antibodies with a labelled anti-human antibody

CC or used for preparing vaccines against dengue virus infection.

SQ Sequence 352 AA;

Query Match 59.2%; Score 45; DB 1; Length 352;

Best Local Similarity 60.0%; Pred. NO. 93;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12

III: I::II

Db 7 SWKNKELKCG 16

RESULT 8

Y05522

ID Y05522 standard; Protein; 1127 AA.

AC Y05522;

DT 05-JUL-1999 (first entry)

DE Dengue virus serotype 2 PR159/S1 viral capsid, prM, E, NS1.

KW Flavivirus; envelope protein; vaccine; infection; diagnosis.

OS Dengue virus serotype 2.

PS Key Location/Qualifiers

FT Protein 1..114

FT /label= Capsid

FT Protein 115..205

FT /label= PreMembrane

FT Protein 206..280

FT /label= Membrane

FT Protein 280..1127

FT /label= Envelope

PN WO9906068-A2.

PD 11-FEB-1999.

PF 27-JUL-1997; U15447.

PR 31-JUL-1997; US-904227.

PA (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.

PI Coller BG, Harada KE, Ivy JM, McDonnell M, Peters ID;

DR WPI; 99-153454/13.

DR N-PSDB; X25114.

DR Recombinant dimeric flaviviral envelope vaccine - comprising a

PT dimeric 80% E protein, useful for protecting against flavivirus,

PT especially dengue virus infections

PS Example 1; Fig 3A-D; 60pp; English.

CC This sequence is composed of the capsid, prM, envelope (E) and NS1

CC proteins of serotype 2 dengue virus DEN-2 strain PR159/S1. A

CC vaccine for protecting against flavivirus infection comprises a

CC dimeric 80% E protein that has been secreted as a recombinant

CC protein from a eukaryotic cell. 80% E indicates a C-terminally

CC truncated flavivirus E protein. The dimeric truncated E is

CC formed: (1) by directly linking 2 tandem copies of 80% E via a

CC flexible tether; (2) via the formation of a leucine zipper domain

CC through the homodimeric association of 2 leucine zipper helices

CC each fused to the C-terminus of an 80% E molecule; or (3) via the

CC formation of a non-covalently associated four-helix bundle domain

CC formed upon association of two helix-turn-helix moieties attached

CC to the C-terminus of an 80% E molecule. Dimeric truncated DEN-2 E

CC proteins are efficiently secreted by recombinant cells, are easier

CC to purify than intracellular proteins, and generate a high titer

CC neutralising antibody response. The method is generally applicable

CC to flaviviruses, in particular dengue viruses such as DEN-2, where

CC 80% E comprises amino acids 1-395 of DEN-2 E. The products can

CC also be used for diagnosis of infection.

CC Sequence 1127 AA;

SQ

Query Match 59.2%; Score 45; DB 1; Length 1127;

```
PI Clements D, Ivy JM, Nakano E;
DR WPI; 97-020938/02.
PT Subunit vaccine against flavivirus infection - contg. recombinant
PT envelope protein in secretable form, used for immunising against
PT flavivirus infection
PS Example 1; Fig 3A-D; 121pp; English.
CC A polypeptide (W09409) comprises the capsid, pre-membrane,
CC envelope and NS1 proteins of dengue virus serotype 2 (DEN-2)
CC variant PR159/S1. A conservative mutation in the envelope protein
CC may be involved in the attenuation of this small-plaque, temp.-
CC sensitive variant. Portions of the envelope protein, esp. domain
CC B, can be expressed in eukaryotic hosts (see also W09410 and
CC W09427-28) transfected with vectors incorporating DEN-2 S1 cDNA
CC (see also T47666). These polypeptides can be used in novel subunit
CC vaccines against viral infection, to raise antibodies useful for
CC passive immunisation, and for diagnosis of infection.
SQ Sequence 1127 AA;

Query Match 59.2%; Score 45; DB 1; Length 1127;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
   |||:|:|
Db 782 SWKNKELKCG 791

RESULT 5
W06591
ID W06591 standard; Protein; 3391 AA.
AC W06591;
DE 12-SEP-1997 (first entry)
DE Polypeptide of attenuated DEN-2 virus, strain 16681, PDK-53.
KW Dengue 2 virus; polypeptide; capsid; prM; M; E; NS1; NS2A; NS2B; NS3;
KW NS4A; NS4B; NS5; PDK-53; quadravalent vaccine; immunity; serotype;
KW chimeric DEN-2/1 virus; chimeric DEN-2/3 virus; chimeric DEN-2/4 virus;
KW dengue fever; fatal dengue haemorrhagic fever; dengue shock syndrome;
KW DHF; DSS.
OS Dengue 2 virus, strain 16681.
FH Key Location/Qualifiers
FT protein 2..114
FT /label= Capsid_protein
FT protein 115..205
FT /label= prM
FT modified_site 183
FT /note= "N-linked glycosylation site, encoded by NAC"
FT protein 206..280
FT /label= M
FT protein 281..775
FT /label= E
FT modified_site 347
FT /note= "N-linked glycosylation site, encoded by NAC"
FT modified_site 433
FT /note= "N-linked glycosylation site, encoded by NAT"
FT protein 776..1127
FT /label= NS1
FT modified_site 905
FT /note= "N-linked glycosylation site, encoded by NAC"
FT modified_site 982
FT /note= "N-linked glycosylation site, encoded by NAT"
FT protein 1128..1345
FT /label= NS2A
FT protein 1346..1475
FT /label= NS2B
FT protein 1476..2093
FT /label= NS3
FT protein 2094..2242
FT /label= NS4A
FT protein 2243..2491
FT /label= NS4B
FT protein 2492..3391
```

```
FT misc_difference 3038 /label= NS5
FT /note= "Encoded by KKA"
PN W09640933-A1.
PD 19-DEC-1996. U09209.
PF 06-JUN-1996; US-483292.
PR (UYMA-) UNIV MAHIDOL AT SALAYA.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Bhamarapavati N, Butrapet S, Chang J, Gubler DJ;
PI Halstead SB, Kinney R, Trent DW;
DR WPI; 97-052330/05.
DR N-PSDB; T49304.
PT PDK-53, a clone of infectious attenuated Dengue 2 virus strain 16681
PT - also chimeric DEN-2/1, DEN-2/3 and DEN-1/4 viruses, used as a
PT quadravalent vaccine for protecting against Dengue virus infection
PT Claim 27; Page 122-136; 261pp; English.
PS This sequence represents the polypeptide from attenuated Dengue 2 virus,
CC strain 16681. The attenuated virus is designated PDK-53. The poly-
CC protein comprises the capsid, prM, M, E, NS1, NS2A, NS2B, NS3, NS4A,
CC NS4B and NS5 proteins. A clone of this wildtype viral sequence, PDK-53,
CC may be used in the production of a quadravalent vaccine which provides
CC immunity against all four serotypes of dengue virus. The vaccine also
CC comprises a chimeric DEN-2/1 virus, a chimeric DEN-2/3 virus, and/or a
CC chimeric DEN-2/4 virus. The new quadravalent vaccines are used to
CC protect against infection by all four serotypes of dengue virus, DEN-1,
CC DEN-2, DEN-3 and DEN-4, which can lead to dengue fever or fatal dengue
CC haemorrhagic fever/dengue shock syndrome (DHF/DSS). Host cells are
CC used to produce the recombinant protein products of the DNA constructs
CC which are used in the vaccines.
SQ Sequence 3391 AA;

Query Match 59.2%; Score 45; DB 1; Length 3391;
Best Local Similarity 60.0%; Pred. No. 91;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12
   |||:|:|
Db 782 SWKNKELKCG 791

RESULT 6
W06590
ID W06590 standard; Protein; 3391 AA.
AC W06590;
DE 11-SEP-1997 (first entry)
DE Polypeptide of DEN-2 virus, strain 16681.
KW Dengue 2 virus; polypeptide; capsid; prM; M; E; NS1; NS2A; NS2B; NS3;
KW NS4A; NS4B; NS5; PDK-53; quadravalent vaccine; immunity; serotype;
KW chimeric DEN-2/1 virus; chimeric DEN-2/3 virus; chimeric DEN-2/4 virus;
KW dengue fever; fatal dengue haemorrhagic fever; dengue shock syndrome;
KW DHF; DSS.
OS Dengue 2 virus, strain 16681.
FH Key Location/Qualifiers
FT protein 2..114
FT /label= Capsid_protein
FT protein 115..205
FT /label= prM
FT modified_site 183
FT /note= "N-linked glycosylation site"
FT protein 206..280
FT /label= M
FT protein 281..775
FT /label= E
FT modified_site 347
FT /note= "N-linked glycosylation site"
FT modified_site 433
FT /note= "N-linked glycosylation site"
FT protein 776..1127
FT /label= NS1
FT modified_site 905
FT /note= "N-linked glycosylation site"
```

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 12 AA;

Query Match 77.0%; Score 58.5; DB 1; Length 12;  
 Best Local Similarity 84.6%; Pred. No. 0.0026;  
 Matches 11; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 COSWKSSEIRCGK 13  
 || |||||:|||||  
 Db 1 CQ-WKSSDIRCGK 12

## RESULT 3

RL3166  
 ID R13166 standard; Protein; 3391 AA.

AC R13166;  
 DT 21-NOV-1991 (first entry)  
 DE Proteins encoded by entire Dengue 2 virus genome.  
 KW dengue virus; detection; consensus sequence; Flavivirus; PCR.  
 OS Dengue virus.

FH Key Location/Qualifiers  
 FT peptide 116..205  
 FT /label= prM  
 FT protein 206..280  
 FT /label= M  
 FT protein 281..775  
 FT /label= E  
 FT protein 776..1127  
 FT /label= NS1  
 FT protein 1128..1345  
 FT /label= NS2A  
 FT protein 1346..1474  
 FT /label= NS2B  
 FT protein 1475..2093  
 FT /label= NS3  
 FT protein 2094..2243  
 FT /label= ns4a  
 FT protein 2244..2492  
 FT /label= NS4B  
 FT protein 2493..3391  
 FT /label= NS5  
 FT modified\_site 183  
 FT /label= N-glycosylated  
 FT modified\_site 347  
 FT /label= N-glycosylated  
 FT modified\_site 433  
 FT /label= N-glycosylated  
 FT modified\_site 905  
 FT /label= N-glycosylated  
 FT modified\_site 982  
 FT /label= N-glycosylated  
 FT modified\_site 1134  
 FT /label= N-glycosylated  
 FT modified\_site 1174  
 FT /label= N-glycosylated  
 FT modified\_site 1329  
 FT /label= N-glycosylated  
 FT modified\_site 1369  
 FT /label= N-glycosylated  
 FT modified\_site 2301  
 FT /label= N-glycosylated  
 FT modified\_site 2305  
 FT /label= N-glycosylated  
 FT modified\_site 2457

FT modified\_site /label= N-glycosylated  
 FT 2485  
 FT /label= N-glycosylated  
 FT modified\_site 2644  
 FT /label= N-glycosylated  
 FT modified\_site 2665  
 FT /label= N-glycosylated  
 FT modified\_site 2704  
 FT /label= N-glycosylated  
 FT modified\_site 2714  
 FT /label= N-glycosylated  
 FT modified\_site 2714  
 FT /label= N-glycosylated  
 PN FR2654113-A.  
 PD 10-MAY-1991.  
 PE 09-NOV-1989; 914724.  
 PR 09-NOV-1989; FR-014724.  
 PA (INSP ) INST PASTEUR.  
 PI Vincent D;  
 DR WPI; 91-225002/31.  
 DR N-PSDB; Q12787.  
 PT Detection and identification of Flaviviridae in biological sample  
 PT - by amplifying consensus sequence then hybridisation opt.  
 PT followed by typing, e.g. sequencing amplified prod.  
 PS Disclosure; Fig 3; 24pp; French.  
 CC The dengue 2 virus is an example of a member of the Flaviviridae  
 CC which can be identified using the probe pair of the invention. A  
 CC species-specific sequence can be amplified using the claimed  
 CC oligonucleotides as primers in a PCR reaction (see Q12788 and  
 CC Q12789). Other viruses which can be identified include Japanese  
 CC encephalitis virus and yellow fever virus. All the dengue 2 virus  
 CC proteins are encoded from an uninterrupted genomic sequence.  
 SQ Sequence 3391 AA;

Query Match 59.2%; Score 45; DB 1; Length 3391;  
 Best Local Similarity 60.0%; Pred. No. 91;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SWKSSEIRCG 12  
 |||:|:|:  
 Db 782 SWKNKELKCG 791

## RESULT 4

W09409  
 ID W09409 standard; Protein; 1127 AA.  
 AC W09409;  
 DT 19-MAY-1997 (first entry)  
 DE Dengue virus serotype 2 PR159/S1 polypeptide.  
 KW DEN-2; flavivirus; envelope protein; immunisation; vaccine.  
 OS Dengue virus serotype 2.  
 FH Key Location/Qualifiers  
 FT region 1..114  
 FT /label= Capsid  
 FT region 115..205  
 FT /label= Pre-membrane  
 FT region 206..280  
 FT /label= Membrane  
 FT region 281..775  
 FT /label= Envelope  
 FT domain 296..395  
 FT /label= Domain-B  
 FT misc\_difference 588  
 FT /note= "amino acid residue 588 (val) is ile  
 FT in wild-type PR159"  
 FT 776..1127  
 FT /label= NS1  
 FT WO9637221-A1.  
 PN 28-NOV-1996.  
 PD 24-MAY-1996; U07627  
 PF 24-MAY-1996; US-448734.  
 PR 07-JUN-1995; US-488807.  
 PR 10-JUL-1995; US-500469.  
 PA (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:49:59 ; Search time 135.78 Seconds  
(without alignments)  
2.268 Million cell updates/sec

Title: US-09-124-280A-29

Perfect score: 76

Sequence: 1 COSWSSEIRCGK 13

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	58.5	77.0	12	1 R61476	Peptide neutralisi
2	58.5	77.0	12	1 W21617	Antibiotic potenti
3	45	59.2	3391	1 R13166	Proteins encoded b
4	45	59.2	1127	1 W09409	Dengue virus serot
5	45	59.2	3391	1 W06591	Polyprotein of att
6	45	59.2	3391	1 W06590	Polyprotein of DEN
7	45	59.2	352	1 W72436	Dengue virus type-
8	45	59.2	1127	1 Y05522	Dengue virus serot
9	41	53.9	715	1 W06593	Amino acid sequenc
10	41	53.9	9	1 W72466	Dengue virus type-
11	40	52.6	3396	1 R43662	DEN1-S275/90 (EGAC
12	40	52.6	108	1 R88719	Human antibody lam
13	40	52.6	798	1 W06592	Amino acid sequenc
14	38	50.0	376	1 R15137	Human serotonin ID
15	38	50.0	398	1 R15138	Human serotonin ID
16	38	50.0	374	1 R25049	5HT1 (serotonin) r
17	38	50.0	622	1 R38888	Sequence encoded b
18	38	50.0	390	1 R43060	Human 5-HT-1B rece
19	38	50.0	41	1 R50685	G-protein coupled
20	38	50.0	327	1 R48708	G-protein coupled
21	38	50.0	327	1 W02680	G-protein coupled
22	38	50.0	41	1 W02877	G-protein coupled
23	38	50.0	319	1 W18658	Fragmented human N
24	37	48.7	110	1 W76130	Human ICR-1.1 V-K
25	37	48.7	110	1 W71260	Humanised murine a
26	37	48.7	110	1 W81454	Humanised antibody
27	37	48.7	110	1 Y00786	Antibody against I
28	37	48.7	75	1 Y13046	Human secreted pro
29	37	48.7	74	1 Y11580	Human 5' EST secre
30	36	47.4	122	1 R84442	Hepatitis GB virus
31	36	47.4	260	1 W41989	Flea serine protea
32	36	47.4	127	1 W76128	Murine ICR-1.1 V-K
33	36	47.4	127	1 W71258	Murine antibody IC
34	36	47.4	130	1 W80717	S. pneumoniae prot
35	36	47.4	127	1 W81452	Murine antibody IC
36	36	47.4	127	1 Y00784	Antibody against I
37	35	46.1	107	1 R47222	Antibody light cha
38	35	46.1	379	1 W33185	Corn barnacle G-pr
39	35	46.1	510	1 W70232	Leishmania antigen
40	35	46.1	159	1 W83952	Polypeptide encode
41	35	46.1	159	1 W83952	U. florida type I
42	34	44.7	129	1 W82707	Variable region of
43	34	44.7	97	1 R07317	VK domain of antib

#### ALIGNMENTS

##### RESULT 1

R61476  
ID R61476 standard; peptide; 12 AA.  
AC R61476;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..10  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M.  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 31; Page 22; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which  
CC include units of formula: (A)n, where A is the cationic amino acid Lys  
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
CC Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides.  
SQ Sequence 12 AA;

Query Match 77.0%; Score 58.5; DB 1; Length 12;

Best Local Similarity 84.6%; Pred. No. 0.0026; Mismatches 0; Indels 1; Gaps 1;

Matches 11; Conservative 1;

QY 1 COSWSSEIRCGK 13

|| |||||

DB 1 CQ-WKSSDIRCGK 12

##### RESULT 2

W21617

ID W21617 standard; peptide; 12 AA.

AC W21617;

DT 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #29.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

OS Synthetic.

FH Key Location/Qualifiers

FT disulfide\_bond 1..10

PN W09638163-A1.

PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;

DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic

PT amino acid sequence - reduces dose of antibiotic required

PS Claim 33; Page 27; 37pp; English.

Tumour Necrosis Fa  
Anti-cancer antibo

**THIS PAGE BLANK (USPTO)**

Query Match 57.8%; Score 37; DB 10; Length 320;  
 Best Local Similarity 55.6%; Pred. No. 77;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLWKYK 9  
 |||:| :|  
 DB 240 KRLQWFRG 248

RESULT 14  
 Q85521 ID Q85521 PRELIMINARY; PRT; 1784 AA.  
 AC Q85521;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)  
 DE GAG-POL PRECURSOR POLYPROTEIN GPR80.  
 OS Feline leukemia virus.  
 OC Viruses; Retrovirdae;  
 OC Mammalian type C retroviruses.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=FELV-FAIDS;  
 RA DONAHUE P.R., HOOVER E.A., BELTZ G.A., RIEDEL N., HIRSCH V.M.,  
 RA OVERBAUGH J., MULLINS J.I.;  
 RT "Strong sequence conservation among horizontally transmissible,  
 RT minimally pathogenic feline leukemia viruses.";  
 RL J. Virol. 62:722-731(1988).  
 DR EMBL: M18247; AAA93092.1; -.  
 DR PFAM; PF01140; gag\_MA; 1.  
 DR PFAM; PF01141; gag\_P12; 1.  
 DR PFAM; PF00552; integrase; 1.  
 DR PFAM; PF00075; rnsaseh; 1.  
 DR PFAM; PF00075; rve; 1.  
 DR PFAM; PF00077; rvp; 1.  
 DR PFAM; PF00078; rvt; 1.  
 DR PFAM; PF00098; zf-CCRC; 1.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 KW Signal; Polyprotein; Hydrolyase; Aspartyl protease.  
 FT SIGNAL 1766 1784 POTENTIAL.  
 FT CHAIN 75 576 POTENTIAL.  
 FT CHAIN 75 576 POTENTIAL.  
 FT CHAIN 202 271 POTENTIAL.  
 FT CHAIN 272 519 POTENTIAL.  
 FT CHAIN 520 576 POTENTIAL.  
 FT CHAIN 577 1784 POTENTIAL.  
 FT CHAIN 577 701 POTENTIAL.  
 FT CHAIN 702 1368 POTENTIAL.  
 FT CHAIN 1369 1784 POTENTIAL.  
 SQ SEQUENCE 1784 AA; 200157 MW; 003D5695 CRC32;

Query Match 57.8%; Score 37; DB 12; Length 1784;  
 Best Local Similarity 60.0%; Pred. No. 4.2e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KRLWKYK 10  
 |||:| :|  
 DB 1401 KRGWYRGK 1410

RESULT 15  
 O89811 ID O89811 PRELIMINARY; PRT; 1786 AA.  
 AC O89811;  
 DT 01-NOV-1998 (TEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)  
 DE GAG-POL PRECURSOR POLYPROTEIN GPR80.  
 OS Feline leukemia virus.

OC Viruses; Retrovirdae; Retroviridae;  
 OC Mammalian type C retroviruses.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RICKARD SUBGROUP A;  
 RX MEDLINE; 98362106.  
 RA CHEN H., BECHTEL M.K., SHI Y., PHIPPS A., MATHES L.E., HAYES K.A.,  
 RA ROY-BURMAN P.;  
 RT "pathogenicity induced by feline leukemia virus, Rickard strain,  
 RT subgroup A plasmid DNA (pFRA).";  
 RL J. Virol. 72:7048-7056(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RICKARD SUBGROUP A;  
 RA CHEN H., ROY-BURMAN P.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF052723; AAC31801.1; -.  
 DR PFAM; PF01140; gag\_MA; 1.  
 DR PFAM; PF01141; gag\_P12; 1.  
 DR PFAM; PF00552; integrase; 1.  
 DR PFAM; PF00075; rnsaseh; 1.  
 DR PFAM; PF00665; rve; 1.  
 DR PFAM; PF00077; rvp; 1.  
 DR PFAM; PF00078; rvt; 1.  
 DR PFAM; PF00098; zf-CCRC; 1.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 KW Polyprotein; Hydrolyase; Aspartyl protease.  
 FT CHAIN 75 576 GAG GENE POLYPROTEIN PR65-GAG.  
 FT CHAIN 75 576 VIRION CORE STRUCTURAL PEPTIDE P15.  
 FT CHAIN 202 271 VIRION CORE STRUCTURAL PEPTIDE P12.  
 FT CHAIN 272 519 VIRION CORE STRUCTURAL PEPTIDE P27.  
 FT CHAIN 520 576 VIRION CORE STRUCTURAL PEPTIDE P10.  
 FT CHAIN 577 1786 POL GENE POLYPROTEIN.  
 FT CHAIN 577 701 PROTEASE.  
 FT CHAIN 702 1370 REVERSE TRANSCRIPTASE.  
 FT CHAIN 1371 1786 ENDONUCLEASE/INTEGRASE.  
 SQ SEQUENCE 1786 AA; 200201 MW; 93530DF1 CRC32;

Query Match 57.8%; Score 37; DB 12; Length 1786;  
 Best Local Similarity 60.0%; Pred. No. 4.2e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KRLWKYK 10  
 |||:| :|  
 DB 1403 KRGWYRGK 1412

Search completed: September 7, 1999, 22:47:38  
 Job time: 7958 sec

RA CHERNOV A.V., VOLLMEYER P., WALTER J., TRAUTNER T.A.;  
 RT "Masc2, a CS-DNA-methyltransferase from Ascombolus immersus with  
 RL similarity to methyltransferases of higher organisms.";  
 RL Biol. Chem. 378:1467-1473(1997).  
 DR EMBL: AF030976; AAC03766.1; -;  
 DR PFAM: PF01426; BAH; 1.  
 DR PFAM: PF00145; DNA\_methylase; 3.  
 KW Transferase; Methyltransferase.  
 SQ SEQUENCE 1356 AA; 153536 MW; AAA066BD CRC32;

Query Match 57.8%; Score 37; DB 3; Length 1356;  
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYKG 9  
 |||||: :|  
 Db 387 KRLKWKYKG 395

RESULT 10  
 O59678 PRELIMINARY; PRT; 166 AA.  
 ID O59678 PRELIMINARY; PRT; 166 AA.  
 AC O59678; 1999 (TrEMBLrel. 09, Created)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)  
 DE HYPOTHETICAL 19.0 KD PROTEIN C29A3.16 IN CHROMOSOME II.  
 GN SPBC29A3.16.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKAERT G.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: TO YEAST YOR294W AND HUMAN KIA0112.  
 DR EMBL: AL022299; CAA18393.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 166 AA; 19044 MW; 7445291A CRC32;

Query Match 57.8%; Score 37; DB 3; Length 166;  
 Best Local Similarity 85.7%; Pred. No. 40;  
 Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 4 KWKYKGK 10  
 |||||  
 Db 123 KWKYKGK 129

RESULT 11  
 O42803 PRELIMINARY; PRT; 1336 AA.  
 ID O42803  
 AC O42803;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE DNA-CS-METHYLTRANSFERASE (EC 2.1.1.37) (DNA  
 DE (CYTOSINE-5)-METHYLTRANSFERASE).  
 GN MASC2.  
 OS Ascombolus immersus.  
 OC Plasmid PCG92.  
 OC Eukaryota; Fungi; Ascomycota; Euscomycetes; Pezizales; Ascombolaceae;  
 OC Ascombolus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RN42.  
 RX MEDLINE; 98182583.  
 RA GOYON C.;  
 RT "Sequence and gene content in 35 kb genomic clone mapping in the

RT human Xq27.1 region.";  
 RL Seq. 8:1-4(1997).  
 CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA -  
 CC S-ADENOSYL-L-HOMOCYSTEINE + DNA CONTAINING 5-METHYLCYTOSINE.  
 DR EMBL: Z36933; CAB09661.1; -;  
 DR PFAM: PF01426; BAH; 1.  
 DR PFAM: PF00145; DNA\_methylase; 3.  
 KW Transferase; Methyltransferase; Plasmid.  
 SQ SEQUENCE 1336 AA; 151125 MW; A2042DCF CRC32;

Query Match 57.8%; Score 37; DB 3; Length 1336;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYKG 9  
 |||||: :|  
 Db 387 KRLKWKYKG 395

RESULT 12  
 O93845 PRELIMINARY; PRT; 2207 AA.  
 ID O93845 PRELIMINARY; PRT; 2207 AA.  
 AC O93845;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE DNA POLYMERASE EPSILON HOMOLOG.  
 GN NIMP.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Euscomycetes; Plecomycetes;  
 OC Eurotiales; Trichocomaceae; Emericella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA JAMES S.W., CRAWFORD G.E., WEXLER A.N.;  
 RL "nimp DNA Pol Epsilon gene of Aspergillus nidulans.";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF019254; AAD01637.1; -;  
 SQ SEQUENCE 2207 AA; 252648 MW; 14A06363 CRC32;

Query Match 57.8%; Score 37; DB 3; Length 2207;  
 Best Local Similarity 45.5%; Pred. No. 5.1e+02;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KRLKWKYKGK 11  
 :||| |::|:  
 Db 648 RRLPWAWRGEF 658

RESULT 13  
 O92UX8 PRELIMINARY; PRT; 320 AA.  
 ID O92UX8 PRELIMINARY; PRT; 320 AA.  
 AC O92UX8;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE F15K20.13 PROTEIN.  
 GN F15K20.13.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,  
 RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;  
 RL "Arabidopsis thaliana chromosome II BAC F15K20 genomic sequence.";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC005824; AAC73023.1; -;  
 SQ SEQUENCE 320 AA; 36096 MW; C0DC45B6 CRC32;

RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
DR EMBL; 249907; CAA90086.1; -.  
SQ SEQUENCE 454 AA; 52667 MW; 25F66E2C CRC32;  
  
Query Match 60.9%; Score 39; DB 5; Length 454;  
Best Local Similarity 54.5%; Pred. No. 51;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KRLWKYKGF 11  
:|:|:|:|:  
DB 103 ERKWKDKGEF 113  
  
RESULT 6  
O93287 PRELIMINARY; PRT; 276 AA.  
AC O93287;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE HOMEBOX TRANSCRIPTION FACTOR.  
GN HOXD-11.  
OS Ambystoma mexicanum (Axolotl).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
OC Batrachia; Caudata; Salamandroidea; Ambystomidae; Ambystoma.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE; 95324399.  
RX GARDINER D.M., BLUMBERG B., KOMINE Y., BRYANT S.V.;  
RT "Regulation of Hoxa expression in developing and regenerating axolotl  
RT limbs.";  
RL Development 121:1731-1741(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98370927.  
RA TOROK M.A., GARDINER D.M., SHUBIN N.H., BRYANT S.V.;  
RT "Expression of Hoxd genes in developing and regenerating axolotl  
RT limbs.";  
RL Dev. Biol. 200:225-233(1998).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL; AF031481; AAC34744.1; -.  
DR PFAM; PF00046; homeobox; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
KW Homeobox; DNA-binding; Nuclear protein.  
SQ SEQUENCE 276 AA; 31673 MW; 5C829680 CRC32;  
  
Query Match 60.9%; Score 39; DB 13; Length 276;  
Best Local Similarity 45.5%; Pred. No. 31;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KRLWKYKGF 11  
:|:|:|:|:  
DB 68 ERKWKYKGF 78  
  
RESULT 7  
O27914 PRELIMINARY; PRT; 168 AA.  
AC O27914;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
DE CONSERVED PROTEIN.  
GN MTH1892.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanobacterium.

RA STRAIN-DELTA H;  
RX MEDLINE; 98037514.  
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
RA ALREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,  
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,  
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,  
RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,  
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
RT deltaH: functional analysis and comparative genomics.";  
RL J. Bacteriol. 179:7135-7155(1997).  
DR EMBL; AE000940; AAB86352.1; -.  
SQ SEQUENCE 168 AA; 18891 MW; DFB942B6 CRC32;  
  
Query Match 59.4%; Score 38; DB 1; Length 168;  
Best Local Similarity 55.6%; Pred. No. 28;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 LKWKYKGF 11  
|:|:|:|:  
DB 96 LSKWFRGEF 104  
  
RESULT 8  
Q20418 PRELIMINARY; PRT; 754 AA.  
AC Q20418;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
DE F44C4.1 PROTEIN.  
GN F44C4.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA SIMS M.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; 249910; CAA90124.1; -.  
DR EMBL; 270034; CAA90124.1; JOINED.  
SQ SEQUENCE 754 AA; 87875 MW; 3739D9F7 CRC32;  
  
Query Match 59.4%; Score 38; DB 5; Length 754;  
Best Local Similarity 70.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 RLWKYKGF 11  
|||:|:|:  
DB 720 RLKWLQKGF 729  
  
RESULT 9  
O42731 PRELIMINARY; PRT; 1356 AA.  
AC O42731;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE C5-DNA-METHYLTRANSFERASE.  
GN MASC2.  
OS Ascobolus immersus.  
OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Pezizales; Ascombolaceae;  
OC Ascobolus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RM42;  
RX MEDLINE; 98121243.



```

RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUOT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL: AE001128: AAC66568.1; -.
DR TIGR: BB0170; -.
KW Hypothetical protein.
SQ SEQUENCE 682 AA; 78474 MW; 4233033A CRC32;

Query Match 60.9%; Score 39; DB 2; Length 682;
Best Local Similarity 54.5%; Pred. No. 76;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KRLKWKYKGF 11
Db 369 KRLKVEYSGEF 379
I:|||||I:I

RESULT 3
O33906 PRELIMINARY; PRT; 2004 AA.
AC O33906;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 216.9 KD PROTEIN.
OS Shewanella sp. SCRC-2738.
OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
OC Shewanella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRC-2738;
RX MEDLINE: 97419510.
RA TAKEYANA H., TAKEDA D., YAZAWA K., YAMADA A., MATSUNAGA T.;
RT "Expression of the eicosapentaenoic acid synthesis gene cluster from
RT Shewanella sp. in a transgenic marine cyanobacterium, Synechococcus
RT sp.";
RL Microbiology 143:0-0(0).
DR EMBL: U73935; AAB81125.1; -.
DR PFAM: PF00109; ketoacyl-synt; 1.
DR PFAM: PF01377; Thioester_dehyd; 2.
KW Hypothetical protein.
SQ SEQUENCE 2004 AA; 216919 MW; 19AABB50 CRC32;

Query Match 60.9%; Score 39; DB 2; Length 2004;
Best Local Similarity 62.5%; Pred. No. 2,2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LKWKYKGF 10
Db 1944 IKWKYRGQ 1951
I:||||I:

RESULT 4
Q05782 PRELIMINARY; PRT; 273 AA.
AC Q05782;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE SIMILAR TO S. CEREVISIAE HYPOTHETICAL PROTEIN IN QRI5 5' REGION.
GN L8167.18.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RX MEDLINE: 97313267.
RA JOHNSTON M., HILLIER L., RILES L., ALBERMANN K., ANDRE B., ANSORGE W.,
RA BENES V., BRUCKNER M., DELIUS H., DUBOIS E., DUSTERHOFT A.,
RA ENTIAN K.D., FLOETH M., GOFFEAU A., HEBLING U., HEUMANN K.,
RA HEUSS-NEITZEL D., HILBERT H., HILGER F., KLEINE K., KOTTER P.,
RA LOUIS E.J., MESSENGUY F., MEWES H.W., MOSGA T., MOSTL D.,
RA MULLER-AUER S., NENTWICH U., OBERMAIER B., PIRAVANDI E., POHL T.M.,
RA PORTETELLE D., PURNELLE B., RECHMANN S., RIEGER M., RINKE M., ROSE M.,
RA SCHARFE M., SCHERENS B., SCHOLLER P., SCHWAGER C., SCHWARZ S.,
RA UNDERWOOD A.P., URRESTARAZU L.A., VANDENBOL M., VERHASSLT P.,
RA VIERENDELS F., VOET M., VOLCKAERT G., VOSS H., WAMBUIT R., WEDLER E.,
RA WEDLER H., ZIMMERMANN F.K., ZOLLNER A., HANI J., HOEISEL J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XII.";
RL Nature 387:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RA PAULEY A.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RA WATERSTON R.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RA CHERRY J.M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U14913; AAB67439.1; -.
SQ SEQUENCE 273 AA; 32391 MW; 9EF66719 CRC32;

Query Match 60.9%; Score 39; DB 3; Length 273;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 2 RLKWKYKGF 11
Db 167 RLKWEYKKNY 176
I:||||I:

RESULT 5
Q17511 PRELIMINARY; PRT; 454 AA.
AC Q17511;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE B0491.4 PROTEIN.
GN B0491.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA SULSTON J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:37 ; Search time 160.2 Seconds  
(without alignments)  
4.226 Million cell updates/sec

Title: US-09-124-280A-28

Perfect score: 64

Sequence: 1 KRLWKYKGRF 11

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10.\*

- 1: sp-archaea.\*
- 2: sp-bacteria.\*
- 3: sp-fungi.\*
- 4: sp-human.\*
- 5: sp-invertebrate.\*
- 6: sp-mammal.\*
- 7: sp-mhc.\*
- 8: sp-organelle.\*
- 9: sp-phage.\*
- 10: sp-plant.\*
- 11: sp-rodent.\*
- 12: sp-virus.\*
- 13: sp-vertebrate.\*
- 14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	62.5	485	8	078396 brachyloma
2	39	60.9	682	2	051192
3	39	60.9	2004	3	033906 shewanella
4	39	60.9	273	3	Q05782
5	39	60.9	454	5	Q17511
6	39	60.9	276	13	Q93287
7	38	59.4	168	1	Q27914
8	38	59.4	754	5	Q20418
9	37	57.8	1356	3	Q42731
10	37	57.8	166	3	Q59678
11	37	57.8	1336	3	Q42803
12	37	57.8	2207	3	Q93845
13	37	57.8	320	10	Q9ZUX8
14	37	57.8	1784	12	Q85521
15	37	57.8	1786	12	Q89811
16	36	56.2	957	1	Q26426
17	36	56.2	438	1	Q58314
18	36	56.2	295	1	Q58667
19	36	56.2	758	1	Q58981
20	36	56.2	445	1	Q59214
21	36	56.2	480	2	Q96634
22	36	56.2	398	2	Q76168
23	36	56.2	253	2	Q9ZDQ4
24	36	56.2	910	3	P87211
25	36	56.2	536	4	Q13529
26	36	56.2	569	5	Q18211
27	36	56.2	815	5	Q18422
28	36	56.2	345	8	Q78881
29	36	56.2	773	12	Q39274

30 36 56.2 362 12 Q69261 equine herp  
31 35.5 55.5 610 2 Q32582 escherichia  
32 35 54.7 558 1 O58315 pyrococcus  
33 35 54.7 226 2 Q06073 bacillus me  
34 35 54.7 149 2 O50557 actinobacil  
35 35 54.7 595 2 O67249 aquifex aeo  
36 35 54.7 329 2 O30601 bacillus su  
37 35 54.7 168 2 P73576 synechocyst  
38 35 54.7 337 2 P73759 synechocyst  
39 35 54.7 301 2 Q9ZGM3 leptospira  
40 35 54.7 1246 3 Q12276 saccharomyc  
41 35 54.7 681 3 P87159 schizosacch  
42 35 54.7 394 3 O94708 schizosacch  
43 35 54.7 872 5 P90523 dictyosteli  
44 35 54.7 776 5 O45269 caenorhabdi  
45 35 54.7 735 6 O18817 oryctolagus

#### ALIGNMENTS

RESULT 1

O78396  
ID O78396 PRELIMINARY; PRT; 485 AA.  
AC O78396;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE MATURASE (FRAGMENT).  
GN MATK.  
OS Brachyloma daphnoides.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Asterales; Ericales; Epacridaceae; Brachyloma.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KRON K.A., FULLER R., CRAVY D.M., GADEK P.A., QUINN C.J.;  
RT "Phylogenetic relationships of epacrids and vaccinioids (Ericaceae s.l.) based on matk sequence data."  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF015633; AAC62165.1; -  
DR PFAM; PF01348; Intron\_maturas2; 1.  
KW Chloroplast.  
FT NON\_TER 485 485  
SQ SEQUENCE 485 AA; 57971 MW; 4795DDF8 CRC32;

Query Match 62.5%; Score 40; DB 8; Length 485;  
Best Local Similarity 63.6%; Pred. No. 37;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KRLWKYKGRF 11

Db 259 KDLKWKVEKGRPF 269

RESULT 2

O51192  
ID O51192 PRELIMINARY; PRT; 682 AA.  
AC O51192;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE HYPOTHETICAL 78.5 KD PROTEIN.  
GN BB0170.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35210 / B31;  
RX MEDLINE; 98065943.  
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,

**THIS PAGE BLANK (USPTO)**

Db 184 LTWRYKGN 192

Search completed: September 7, 1999, 23:59:09  
Job time: 521 sec

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M81249; G211937; -;  
 DR EMBL; D10288; G222805; -;  
 DR PIR; B37914; B37914.  
 DR PIR; S14799; S14799.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEBOX\_2; 1.  
 DR PFAM; PF00046; homeobox; 1.  
 DR HSP; P02835; IFT2.  
 DR TRANSFAC; T01759; -;  
 KW HOMEBOX; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN;  
 KW TRANSCRIPTION REGULATION.  
 FT DNA\_BIND 208 267 HOMEBOX.  
 FT CONFLICT 181 200 VITASGAASPPGEAVERKSN -> SSSQPSAVVFCINQHFL  
 FT SLI (IN REF. 2).  
 SQ SEQUENCE 280 AA; 31923 MW; 016DC9ED CRC32;

Query Match 57.8%; Score 37; DB 1; Length 280;  
 Best Local Similarity 45.5%; Pred. No. 16;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 KRLKWKYKGF 11  
 :||:|:|:  
 DB 68 ERGKWRGYSY 78

RESULT 14  
 NU2M\_DASNO STANDARD; PRT; 347 AA.  
 AC O21326;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).  
 GN MTND2 OR ND2 OR NADH2.  
 OS DASYRUS NOVEMINCTUS (NINE-BANDED ARMADILLO).  
 OC MITOCHONDRION.  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC EDENTATA; DASYPODIDAE; DASYRUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97357423.  
 RA ARNASON U., GULLBERG A., JANKE A.;  
 RT "Phylogenetic analyses of mitochondrial DNA suggest a sister group  
 RT relationship between Xenarthra (Edentata) and Ferungulates.";  
 RL MOL. BIOL. EVOL. 14:762-768(1997).  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL  
 CC INNER MEMBRANE.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Y11832; E311017; -;  
 DR PFAM; PF00361; oxidored\_q1; 1.  
 KW OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION; TRANSMEMBRANE.  
 SQ SEQUENCE 347 AA; 39329 MW; 383847A5 CRC32;

Query Match 57.8%; Score 37; DB 1; Length 347;  
 Best Local Similarity 71.4%; Pred. No. 20;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLKWKYK 8  
 :|||:  
 DB 312 KMKWKYK 318

RESULT 15  
 PSAB\_YERPE STANDARD; PRT; 273 AA.  
 ID PSAB\_YERPE  
 AC P31523; Q56979;  
 DT 01-JUL-1993 (REL. 26, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE CHAPERONE PROTEIN PSAB PRECURSOR.  
 GN PSAB.  
 OS YERSINIA PESTIS, AND YERSINIA PSEUDOTUBERCULOSIS.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 OC YERSINIA.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX SPECIES=Y. PESTIS;  
 RX MEDLINE; 93302507.  
 RA LINDLER L.E., TALL B.D.;  
 RT "Yersinia pestis ph 6 antigen forms fimbriae and is induced by  
 RT intracellular association with macrophages.";  
 RL MOL. MICROBIOL. 8:311-324(1993).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX SPECIES=Y. PESTIS; STRAIN-EV76;  
 RA CHEREPAVOV P.A.;  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX SPECIES=Y. PSEUDOTUBERCULOSIS; STRAIN=YP111;  
 RX MEDLINE; 96294755  
 RA YANG Y., MERRIAM J.J., MUELLER J., ISBERG R.R.;  
 RT "The psa locus is responsible for thermoinducible binding of Yersinia  
 RT pseudotuberculosis to cultured cells.";  
 RL INFECT. IMMUN. 64:2483-2489(1996).  
 CC -1- FUNCTION: REQUIRED FOR THE BIOGENESIS OF THE PH6 ANTIGEN.  
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).  
 CC -1- SIMILARITY: STRONG, TO OTHER PERIPLASMIC PILUS CHAPERONES.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN GENE SUPERFAMILY.  
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-23 IS THE  
 CC INITIATOR.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M86713; G155489; ALT\_INIT.  
 DR EMBL; X97759; E242706; -;  
 DR EMBL; L76301; G1197044; -;  
 DR PIR; S27748; S27748.  
 DR PROSITE; PS00635; PILLI\_CHAPERONE; 1.  
 DR PFAM; PF00345; Pili\_assembly; 1.  
 KW CHAPERONE; PERIPLASMIC; SIGNAL; IMMUNOGLOBULIN FOLD.  
 FT SIGNAL 1 31 POTENTIAL.  
 FT CHAIN 32 273 CHAPERONE PROTEIN PSAB.  
 FT DISULFID 128 163 POTENTIAL.  
 FT CONFLICT 252 252 E -> D (IN REF. 2).  
 SQ SEQUENCE 273 AA; 30648 MW; 9B4E5D43 CRC32;

Query Match 57.8%; Score 37; DB 1; Length 273;  
 Best Local Similarity 55.6%; Pred. No. 16;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 LKWKYKGF 11  
 :||:|:|:

DE 60S RIBOSOMAL PROTEIN L10.  
GN REL10.  
OS EUGLENA GRACILIS.  
OC EUKARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.  
RN [1]  
RC STRAIN=2;  
RA DERUERE J., SCHANTZ M.L., SCHANTZ R.;  
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: BELONGS TO THE L10E FAMILY OF RIBOSOMAL PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X94087; E213753; .  
DR PROSITE; PS01257; RIBOSOMAL\_L10E; 1.  
DR PFAM; PF00826; Ribosomal\_L10e; 1.  
KW RIBOSOMAL PROTEIN.  
SQ SEQUENCE 215 AA; 24642 MW; CBF0FC23 CRC32;  
  
Query Match 59.4%; Score 38; DB 1; Length 215;  
Best Local Similarity 66.7%; Pred. No. 8.2;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 LKWKYKGF 11  
Db 182 VKWKEGKF 190  
  
RESULT 12  
UDAL\_RAT STANDARD; PRT; 527 AA.  
AC F36510;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE UDP-GLUCURONOSYLTRANSFERASE 2A1 PRECURSOR, MICROSOMAL (EC 2.4.1.17)  
DE (UDPCT) (UGT-OLF).  
GN UGT2A1 OR UGT2A-1.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91156050.  
RA LAZARD D., ZUPKO K., PORIA Y., NEF P., LAZAROVITS J., HORN S.,  
RA KHEN M., LANCET D.;  
RT "Odorant signal termination by olfactory UDP glucuronosyl  
RT transferase";  
RL NATURE 349:790-793(1991).  
CC -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND  
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND  
CC ENDOGENOUS COMPOUNDS. THIS ISOFORM IS ACTIVE ON ODORANTS AND  
CC SEEMS TO BE INVOLVED IN OLFACTION; IT COULD HELP CLEAR LIPOPHILIC  
CC ODORANT MOLECULES FROM THE SENSORY EPITHELIUM.  
CC -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR  
CC BETA-D-GLUCURONOSIDE.  
CC -!- TISSUE SPECIFICITY: OLFACTORY EPITHELIUM.  
CC -!- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL: X57565; .; NOT\_ANNOTATED\_CDS.  
DR PROSITE; PS00375; UDPGT; 1.  
DR PFAM; PF00201; UDPGT; 1.  
KW TRANSFERASE; GLYCOSYLTRANSFERASE; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL;  
KW MULTIGENE FAMILY; MICROSOME; OLFACTION.  
FT SIGNAL 1 20  
FT CHAIN 21 527  
FT TRANSMEM 491 507  
FT CARBOHYD 49 49  
FT CARBOHYD 313 313  
SQ SEQUENCE 527 AA; 59915 MW; F81D3031 CRC32;  
  
Query Match 59.4%; Score 38; DB 1; Length 527;  
Best Local Similarity 50.0%; Pred. No. 21;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KRLWKYKGF 10  
Db 331 OKVLMRYKGF 340  
  
RESULT 13  
HXDB\_CHICK STANDARD; PRT; 280 AA.  
AC P24342; Q90780;  
DT 01-MAR-1992 (REL. 21, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HOMEBOX PROTEIN HOX-D11 (CHOX-4.6) (GHOX-4.6) (CHOX-4E).  
GN HOXD11 OR CHOX-4.6.  
OS GALLUS GALLUS (CHICKEN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92173259.  
RA REGINA B., COELHO C.N., KOSHER R.A., UPHOLT W.B.;  
RT "The pattern of expression of the chicken homolog of HOX11 in the  
RT developing limb suggests a possible role in the ectodermal inhibition  
RT of chondrogenesis";  
RL DEV. DYN. 193:92-101(1992).  
RN [2]  
RP SEQUENCE OF 181-280 FROM N.A.  
RA NOHNO T.;  
RL SUBMITTED (FEB-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE OF 208-268 FROM N.A.  
RX MEDLINE; 91168267.  
RA NOHNO T., NOJI S., KOYAMA E., OHYAMA K., MYOKAI F., KUROIWA A.,  
RA SAITO T., TANIGUCHI S.;  
RT "Involvement of the Chox-4 chicken homeobox genes in determination of  
RT anteroposterior axial polarity during limb development";  
RL CELL 64:1197-1205(1991).  
RN [4]  
RP SEQUENCE OF 208-267 FROM N.A.  
RX MEDLINE; 91204053.  
RA IZPISUA-BELMONT J.-C., TICKLE C., DOLLE P., WOLPERT L., DUBOULE D.;  
RT "Expression of the homeobox Hox-4 genes and the specification of  
RT position in chick wing development";  
RL NATURE 350:585-589(1991).  
CC -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF  
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH  
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- DEVELOPMENTAL STAGE: COORDINATELY EXPRESSED IN PARTIALLY  
CC OVERLAPPING DOMAINS DURING WING DEVELOPMENT.  
CC -!- SIMILARITY: BELONGS TO THE ARD-B FAMILY OF HOMEBOX PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92397593.  
 RA SIMON M., DELLA SETA F., SOR F., FAYE G.;  
 RT "Analysis of the MSS51 region on chromosome XII of *Saccharomyces*  
 cerevisiae";  
 RL YEAST 8:559-567(1992).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; J01487; -; NOT\_ANNOTATED\_CDS.  
 CC DR EMBL; S43721; G25247; -;  
 CC DR PIR; S25342; S25342.  
 KW HYPOTHETICAL PROTEIN.  
 FT NON\_TER 1  
 SQ SEQUENCE 106 AA; 12578 MW; 19E7F9B9 CRC32;

Query Match 60.9%; Score 39; DB 1; Length 106;  
 Best Local Similarity 60.08; Pred. No. 2.7;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RLKWKYKGF 11  
 |||||:  
 DB 90 RLKWEYKKNY 99

RESULT 10  
 CD47\_HUMAN  
 ID CD47\_HUMAN STANDARD; PRT; 323 AA.  
 AC Q08722;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE LEUKOCYTE SURFACE ANTIGEN CD47 PRECURSOR (ANTIGENIC SURFACE  
 DE DETERMINANT PROTEIN OA3) (INTEGRIN ASSOCIATED PROTEIN) (IAP) (MER6).  
 GN CD47.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OVARY.  
 RX MEDLINE; 93007897.  
 RA CAMPBELL I.G., FREEMONT P.S., FOULKES W., TROWSDALE J.;  
 RT "An ovarian tumor marker with homology to vaccinia virus contains an  
 RT IgV-like region and multiple transmembrane domains.";  
 RL CANCER RES. 52:5416-5420(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94012992.  
 RA LINDBERG F.P., GRESHAM H.D., SCHWARZ E., BROWN E.J.;  
 RT "Molecular cloning of integrin-associated protein: an immunoglobulin  
 RT family member with multiple membrane-spanning domains implicated in  
 RT alpha v beta 3-dependent ligand binding.";  
 RL J. CELL BIOL. 123:485-496(1993).  
 RN [3]  
 RP PARTIAL SEQUENCE, AND IDENTIFICATION AS CD47.  
 RC TISSUE-ERYTHROCYTE;  
 RX MEDLINE; 95091687.  
 RA MANNING W.J., HOLMES C.H., ANSTEE D.J., SPRING F.A., TANNER M.J.A.;  
 RT "Isolation and characterization of CD47 glycoprotein: a multispanning  
 RT membrane protein which is the same as integrin-associated protein  
 RT (IAP) and the ovarian tumour marker OA3.";  
 RL BIOCHEM. J. 304:525-530(1994).

CC -----  
 CC -!- FUNCTION: MAY PLAY A ROLE IN MEMBRANE TRANSPORT AND/OR SIGNAL  
 CC TRANSDUCTION. MAY BE INVOLVED IN MEMBRANE PERMEABILITY CHANGES  
 CC INDUCED FOLLOWING VIRUS INFECTION. HAS A ROLE IN CELL ADHESION IN  
 CC NON-ERYTHROID CELLS.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: VERY BROADLY DISTRIBUTED ON NORMAL ADULT  
 CC TISSUES, AS WELL AS OVARIAN TUMORS, BEING ESPECIALLY ABUNDANT IN  
 CC SOME EPITHELIA AND THE BRAIN.  
 CC -!- PTM: MAY HAVE TWO INTRAMOLECULAR DISULFIDE BONDS.  
 CC -!- ALTERNATIVE PRODUCTS: FOUR FORMS OF THIS PROTEIN ARE PRODUCED BY  
 CC ALTERNATIVE SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF OA3-323.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
 CC ONE V-LIKE DOMAIN.  
 CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD47 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd47.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; X69398; G396176; -;  
 CC DR EMBL; Z25521; G396705; -;  
 CC DR MIM; 601028; -;  
 CC DR PFM; PF00047; ig. 1.  
 KW ANTIGEN; TRANSMEMBRANE; GLYCOPROTEIN; ALTERNATIVE SPLICING; SIGNAL.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 323 LEUKOCYTE SURFACE ANTIGEN CD47.  
 FT MOD\_RES 19 19 BLOCKED (PROBABLE).  
 FT DOMAIN 19 139 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 140 162 POTENTIAL.  
 FT DOMAIN 163 176 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 177 199 POTENTIAL.  
 FT DOMAIN 200 210 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 211 231 POTENTIAL.  
 FT DOMAIN 232 247 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 248 264 POTENTIAL.  
 FT DOMAIN 265 267 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 268 292 POTENTIAL.  
 FT DOMAIN 293 323 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 19 76 IG-LIKE V-TYPE DOMAIN.  
 FT CARBOHYD 23 23 POTENTIAL.  
 FT CARBOHYD 34 34 POTENTIAL.  
 FT CARBOHYD 50 50 POTENTIAL.  
 FT CARBOHYD 73 73 POTENTIAL.  
 FT CARBOHYD 111 111 POTENTIAL.  
 FT CARBOHYD 206 206 MISSING (IN OA3-293).  
 FT VARSPIC 293 323 KA -> NN (IN OA3-305).  
 FT VARSPIC 304 305 MISSING (IN OA3-305).  
 FT VARSPIC 306 323 MISSING (IN OA3-312).  
 FT VARSPIC 312 323 MISSING (IN OA3-312).  
 SQ SEQUENCE 323 AA; 35213 MW; AA2718DB CRC32;

Query Match 59.4%; Score 38; DB 1; Length 323;  
 Best Local Similarity 62.5%; Pred. No. 12;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LKWKYKGF 10  
 |||||:  
 DB 56 VKWFKGR 63

RESULT 11  
 RL10\_EUGGR  
 ID RL10\_EUGGR STANDARD; PRT; 215 AA.  
 AC Q39724;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DR HSP; P02835; 1FTZ.  
DR TRANSFAC; T01761; .  
KW HOMBOX; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN;  
KW TRANSCRIPTION REGULATION.  
FT DNA\_BIND 205 264. HOMEBOX.  
SQ SEQUENCE 277 AA; 31782 MW; A3CBF4BD CRC32;  
  
Query Match 62.5%; Score 40; DB 1; Length 277;  
Best Local Similarity 45.5%; Pred. No. 4.8;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KRLWKYKGF 11  
: : ||:|:|:  
DB 68 RRSKWYRGYS 78  
  
RESULT 8  
LAR\_DROME STANDARD; PRT: 2029 AA.  
ID LAR\_DROME  
AC P16621;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE PROTEIN-TYROSINE PHOSPHATASE DLAR PRECURSOR (EC 3.1.3.48) (PROTEIN-TYROSINE-PHOSPHATE PHOSPHOHYDROLASE).  
GN LAR.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE: 90046860.  
RX STREULI M., KRUEGER N.X., TSAI A.Y.M., SAITO H.;  
RA "A family of receptor-linked protein tyrosine phosphatases in humans and Drosophila.";  
RT PROC. NATL. ACAD. SCI. U.S.A. 86:8698-8702(1989).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN-CANTON-S;  
RX MEDLINE: 96178473.  
RA KRUEGER N.X., VAN VACTOR D., WAN H.I., GELBART W.M., GOODMAN C.S., SAITO H.;  
RT "The transmembrane tyrosine phosphatase DLAR controls motor axon guidance in Drosophila.";  
RL CELL 84:611-622(1996).  
CC -!- FUNCTION: IT IS POSSIBLE THAT DLAR IS A CELL ADHESION RECEPTOR. IT POSSESSES AN INTRINSIC PROTEIN TYROSINE PHOSPHATASE ACTIVITY (PTPASE). IT CONTROLS MOTOR AXON GUIDANCE.  
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED IN A SUBSET OF AXONS AND PIONEER NEURONS IN THE EMBRYO.  
CC -!- SIMILARITY: EXTRACELLULAR REGION TYPIC OF A CAM FAMILY (3 IG-LIKE DOMAINS FOLLOWED BY 9 FIBRONECTIN TYPE III-LIKE DOMAINS), AND A CYTOPLASMIC REGION COMPOSED OF TWO PTPASE DOMAINS.  
CC -!- SIMILARITY: TO DPTP AND TO LAR.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M27700; G157812; .  
DR EMBL; U36857; G1209648; .  
DR EMBL; U36849; G1209648; JOINED.  
DR EMBL; U36850; G1209648; JOINED.  
DR EMBL; U36851; G1209648; JOINED.

DR EMBL; U36852; G1209648; JOINED.  
DR EMBL; U36853; G1209648; JOINED.  
DR EMBL; U36854; G1209648; JOINED.  
DR EMBL; U36855; G1209648; JOINED.  
DR EMBL; U36856; G1209648; JOINED.  
DR PIR; A36182; TDFELK.  
DR FLYBASE; FBgn0000464; Lar.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 2.  
DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 2.  
DR PROSITE; PS50055; TYR\_PHOSPHATASE\_PTP; 2.  
DR PFAM; PF00041; fn3; 9.  
DR PFAM; PF00047; 19; 3.  
DR PFAM; PF00102; Y\_phosphatase; 2.  
DR HSDP; P28827; 1RPM.  
KW HYDROLASE; RECEPTOR; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
KW CELL ADHESION; IMMUNOGLOBULIN FOLD; DUPLICATION.  
FT SIGNAL 1 32  
FT CHAIN 33 2029  
FT DOMAIN 33 1377  
FT TRANSMEM 1378 1402  
FT DOMAIN 1403 2029  
FT DOMAIN 50 118  
FT DOMAIN 154 216  
FT DOMAIN 249 308  
FT DOMAIN 320 417  
FT DOMAIN 418 512  
FT DOMAIN 513 607  
FT DOMAIN 608 706  
FT DOMAIN 707 809  
FT DOMAIN 810 906  
FT DOMAIN 907 1007  
FT DOMAIN 1008 1102  
FT DOMAIN 1103 1207  
FT DOMAIN 1492 1738  
FT DOMAIN 1781 2029  
FT ACT\_SITE 1670 1670  
FT ACT\_SITE 1961 1961  
FT DISULFID 57 111  
FT DISULFID 161 209  
FT DISULFID 256 301  
FT CARBOHYD 176 176  
FT CARBOHYD 233 253  
FT CARBOHYD 298 298  
FT CARBOHYD 553 553  
FT CARBOHYD 616 616  
FT CARBOHYD 666 666  
FT CARBOHYD 721 721  
FT CARBOHYD 774 774  
FT CARBOHYD 915 915  
FT CARBOHYD 962 962  
FT CARBOHYD 1183 1183  
FT CARBOHYD 1304 1304  
SQ SEQUENCE 2029 AA; 229027 MW; 6BDF8F93 CRC32;  
  
Query Match 60.9%; Score 39; DB 1; Length 2029;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 RLWKYK 9  
: : ||:|:|:  
DB 339 RLEWYSYK 346  
  
RESULT 9  
YQR5\_YEAST STANDARD; PRT: 106 AA.  
ID YQR5\_YEAST  
AC P32339;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL PROTEIN IN OR15 5'REGION (FRAGMENT).  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).



```

FT TRANSMEM 492 512 POTENTIAL.
FT CARBOHYD 7 7 POTENTIAL.
FT CARBOHYD 20 20 POTENTIAL.
FT CARBOHYD 248 248 POTENTIAL.
FT CARBOHYD 341 341 POTENTIAL.
SQ SEQUENCE 563 AA: 62056 MW: 2ED2DAA7 CRC32:

Query Match 64.1% Score 41; DB 1; Length 563;
Best Local Similarity 87.5%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 LKWYKKGK 10
I | | | | |
Db 508 LWYKKGK 515

RESULT 6
DPOE_YEAST STANDARD; PRT; 2222 AA.
AC P21951:
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA POLYMERASE EPSILON, CATALYTIC SUBUNIT A (EC 2.7.7.7) (DNA
DE POLYMERASE II SUBUNIT A).
GN POL2 OR DUN2 OR YNL262W OR N0825.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMICETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1214-1221.
RX MEDLINE: 90381771.
RA MORRISON A., ARAKI H., CLARK A.B., HAMATAKE R.K., SUGINO A.;
RT "A third essential DNA polymerase in S. cerevisiae.";
RL CELL 62:1143-1151(1990).
[2]
RN SEQUENCE OF 1-2221 FROM N.A.
RC STRAIN-S288C / FV1679;
RX MEDLINE: 96310631.
RA SEN-GUPTA M., LYCK R., FLEIG U., NIEDENTHAL R.K., HEGERMANN J.H.;
RT "The sequence of a 24,152 bp segment from the left arm of chromosome
RT XIV from Saccharomyces cerevisiae between the BNI1 and the POL2
genes.";
RL YEAST 12:505-514(1996).
[3]
RN TEMPERATURE SENSITIVE MUTANTS.
RP MEDLINE: 92164663.
RA ARAKI H., ROPP P.A., JOHNSON A.L., JOHNSTON L.H., MORRISON A.,
RA SUGINO A.;
RT "DNA polymerase II, the probable homolog of mammalian DNA polymerase
RT epsilon, replicates chromosomal DNA in the yeast Saccharomyces
RT cerevisiae.";
RL EMBO J. 11:733-740(1992).
CC -1- FUNCTION: DNA POLYMERASE II PARTICIPATES IN CHROMOSOMAL DNA
CC REPLICATION.
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -1- SUBUNIT: CONSISTS OF FIVE SUBUNITS (200 KD, 80 KD, 34 KD, 30 KD,
CC AND 29 KD).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DOMAIN: THE DNA POLYMERASE ACTIVITY DOMAIN RESIDES IN THE
CC N-TERMINAL HALF OF THE PROTEIN, WHILE THE C-TERMINUS IS NECESSARY
CC FOR COMPLEXING SUBUNITS B AND C.
CC -1- IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES: ALPHA, BETA, GAMMA,
CC DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR DIFFERENT REACTIONS
CC OF DNA SYNTHESIS.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY. HIGH
CC SIMILARITY WITH MAMMALIAN DNA POLYMERASE EPSILON.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: M60416; G171409; -.
CC EMBL: X92494; G1045247; -.
CC EMBL: Z71538; E239648; -.
CC PIR: A36028; A36028.
CC SGD: L0001461; POL2.
CC PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.
CC PFAM: PF00136; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION;
CC TRANSFERASE; ZINC-FINGER; NUCLEAR PROTEIN.
CC KW TRANSFERASE; ZINC-FINGER; NUCLEAR PROTEIN.
FT ZN_FING 2108 2181 POTENTIAL.
FT SIMILAR 349 388 CONSERVED REGION IV.
FT SIMILAR 604 668 CONSERVED REGION II.
FT SIMILAR 768 786 CONSERVED REGION VI.
FT SIMILAR 817 858 CONSERVED REGION III.
FT SIMILAR 871 879 CONSERVED REGION I.
FT SIMILAR 976 992 CONSERVED REGION V.
FT VARIANT 644 644 M -> I (IN POL2-9 TS MUTANT).
FT VARIANT 710 710 P -> S (IN POL2-18 TS MUTANT).
SQ SEQUENCE 2222 AA; 255669 MW; 233EFA0E CRC32;

Query Match 62.5% Score 40; DB 1; Length 2222;
Best Local Similarity 45.5%; Pred. No. 40;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KRLWKYKGF 11
: | | | : | |
Db 679 RKLAWRGEF 689

RESULT 7
HXDB_NOTVI STANDARD; PRT; 277 AA.
AC P31263:
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE HMOEBOX PROTEIN HOX-D11 (NVHox-2).
GN HOXD11 OR NVHox2.
OS NOTOPHTHALMUS VIRIDESCENS (EASTERN NEWT) (TRITURUS VIRIDESCENS).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA;
OC CAUDATA; SALAMANDROIDEA; SALAMANDRIDAE; NOTOPHTHALMUS.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE: 91372146.
RA BROWN R., BROCKES J.P.;
RT "Identification and expression of a regeneration-specific homeobox
RT gene in the newt limb blastema.";
RL DEVELOPMENT 111:489-496(1991).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIVE IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ABD-B FAMILY OF HMOEBOX PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X57674; -. NOT_ANNOTATED_CDS.
CC PIR: A43783; A43783.
CC PROSITE: PS00027; HMOEBOX_1; 1.
CC PROSITE: PS50071; HMOEBOX_2; 1.
CC PFAM: PF00046; homeobox; 1.

```

RC TISSUE=LIVER;  
RX MEDLINE; 90251467.  
RA GRUNDMANN U., NERLICH C., REIN T., ZETTLMEISSL G.;  
RT "Complete cDNA sequence encoding the B subunit of human factor XIII.";  
RL NUCLEIC ACIDS RES. 18:2817-2817(1990).  
RN (5)  
RP VARIANT PHE-450.  
RX MEDLINE; 93313189.  
RA HASHIGUCHI T., SAITO M., MORISHITA E., MATSUDA T., ICHINOSE A.;  
RT "Two genetic defects in a patient with complete deficiency of the b-  
RT subunit for coagulation factor XIII.";  
RL BLOOD 82:145-150(1993).  
CC -1- FUNCTION: THE B CHAIN OF FACTOR XIII IS NOT CATALYTICALLY ACTIVE,  
CC BUT IS THOUGHT TO STABILIZE THE A SUBUNITS AND REGULATE THE RATE  
CC OF TRANSGLUTAMINASE FORMATION BY THROMBIN.  
CC -1- SUBUNIT: TETRAMER OF TWO A CHAINS AND TWO B CHAINS.  
CC -1- DISEASE: A DEFICIENCY IN FAL3 CAN RESULT IN A LIFELONG BLEEDING  
CC TENDENCY, DEFECTIVE WOUND HEALING, AND HABITUAL ABORTION.  
CC -1- SIMILARITY: CONTAINS 10 SUSHI (SCR) REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M64554; G179417; ALT\_SEQ.  
DR EMBL; M14057; G182839; -.  
DR EMBL; X51823; G31505; -.  
DR PIR; A23830; A23830.  
DR PIR; A36397; A36397.  
DR PIR; S09980; S09980.  
DR MIM; 134580; -.  
DR PFAM; PF00084; sushi; 8.  
DR HSSP; P08603; IHFI.  
DR TRANSFERASE; PLASMA; BLOOD COAGULATION; REPEAT; GLYCOPROTEIN; SIGNAL;  
KW SUSHI; DISEASE MUTATION.  
FT SIGNAL 1 20  
FT CHAIN 21 661 FACTOR XIII, B CHAIN.  
FT DOMAIN 24 647 10 X SUSHI (SCR) REPEATS.  
FT REPEAT 24 88 SUSHI 1.  
FT REPEAT 90 147 SUSHI 2.  
FT REPEAT 152 209 SUSHI 3.  
FT REPEAT 212 268 SUSHI 4.  
FT REPEAT 273 328 SUSHI 5.  
FT REPEAT 335 390 SUSHI 6.  
FT REPEAT 395 451 SUSHI 7.  
FT REPEAT 453 516 SUSHI 8.  
FT REPEAT 523 579 SUSHI 9.  
FT REPEAT 581 647 SUSHI 10.  
FT DISULFID 25 76 BY SIMILARITY.  
FT DISULFID 59 87 BY SIMILARITY.  
FT DISULFID 91 135 BY SIMILARITY.  
FT DISULFID 118 146 BY SIMILARITY.  
FT DISULFID 153 197 BY SIMILARITY.  
FT DISULFID 180 208 BY SIMILARITY.  
FT DISULFID 213 255 BY SIMILARITY.  
FT DISULFID 241 267 BY SIMILARITY.  
FT DISULFID 274 316 BY SIMILARITY.  
FT DISULFID 302 327 BY SIMILARITY.  
FT DISULFID 336 378 BY SIMILARITY.  
FT DISULFID 364 389 BY SIMILARITY.  
FT DISULFID 396 439 BY SIMILARITY.  
FT DISULFID 425 450 BY SIMILARITY.  
FT DISULFID 454 505 BY SIMILARITY.  
FT DISULFID 486 515 BY SIMILARITY.  
FT DISULFID 524 567 BY SIMILARITY.  
FT DISULFID 553 578 BY SIMILARITY.  
FT DISULFID 582 636 BY SIMILARITY.  
FT DISULFID 616 646 BY SIMILARITY.  
FT CARBOHYD 162 162 POTENTIAL.

FT CARBOHYD 545 545 POTENTIAL.  
FT SITE 617 619 CELL ATTACHMENT SITE.  
FT VARIANT 450 450 C -> F (IN F13B DEFICIENCY).  
SQ SEQUENCE 661 AA; 75491 MW; FDD245E9 CRC32;  
  
Query Match 65.6%; Score 42; DB 1; Length 661;  
Best Local Similarity 75.0%; Pred. No. 5.3;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 LKWKYK GK 10  
Db 468 MKWKYEGK 475  
  
RESULT 5  
HNM1\_YEAST STANDARD; PRT; 563 AA.  
ID HNM1\_YEAST  
AC F19807;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE CHOLINE TRANSPORT PROTEIN.  
GN HNM1 OR CTRL OR YGL077C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=X2180-1B;  
RX MEDLINE; 90368823.  
RA NIKAWA J.-I., HOSAKA K., TSUKAGOSHI Y., YAMASHITA S.;  
RT "Primary structure of the yeast choline transport gene and regulation  
RT of its expression";  
RL J. BIOL. CHEM. 265:15996-16003(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE; 97435481.  
RA RIEGER M., BRUECKNER M., SCHAEFER M., MUELLER-AUER S.;  
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae  
RT chromosome VII.";  
RL YEAST 13:1077-1090(1997).  
CC -1- FUNCTION: SOLE CHOLINE TRANSPORTER IN YEAST.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- INHIBITION OF ACTIVITY BY INTRACELLULAR CHOLINE.  
CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J05603; G171330; -.  
DR EMBL; 272595; E243965; -.  
DR PIR; S11175; S11175.  
DR SGD; L0000795; HNM1.  
DR PROSITE; PS00218; AMINO\_ACID\_PERMEASE; 1.  
DR PFAM; PF00324; aa\_permeases; 1.  
KW TRANSPORT; AMINO-ACID TRANSPORT; TRANSMEMBRANE.  
FT TRANSMEM 58 78 POTENTIAL.  
FT TRANSMEM 88 108 POTENTIAL.  
FT TRANSMEM 183 203 POTENTIAL.  
FT TRANSMEM 206 226 POTENTIAL.  
FT TRANSMEM 256 276 POTENTIAL.  
FT TRANSMEM 294 314 POTENTIAL.  
FT TRANSMEM 343 363 POTENTIAL.  
FT TRANSMEM 399 417 POTENTIAL.  
FT TRANSMEM 427 445 POTENTIAL.  
FT TRANSMEM 466 486 POTENTIAL.

CC LPS-MEDIATED ACTIVATION OF THE HEMOLYMPH COAGULATION. IT HAS A  
 CC STRONG ANTIBACTERIAL EFFECT ESPECIALLY ON THE GROWTH OF GRAM-  
 CC NEGATIVE BACTERIA.  
 DR PIR: A27819; A27819.  
 KW ANTIBIOTIC. 31 52 K -> N.  
 FT DISULFID 13 13  
 FT VARIANT 101 AA; 11801 MW; 32EB8EF CRC32;  
 SQ SEQUENCE

Query Match 95.3%; Score 61; DB 1; Length 101;  
 Best Local Similarity 90.9%; Pred. No. 0.0004;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLWKYKGF 11  
 :|||||||  
 Db 40 RRLWKYKGF 50

RESULT 3  
 F13B\_MOUSE STANDARD; PRT; 668 AA.  
 AC Q07968:  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE COAGULATION FACTOR XIII B CHAIN PRECURSOR (EC 2.3.2.13) (PROTEIN-  
 DE GLUTAMINE GAMMA-GLUTAMYLTRANSFERASE B CHAIN) (TRANSGLUTAMINASE B  
 DE CHAIN).  
 GN F13B OR CF13B.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-B10 D2/OSN; TISSUE=LIVER;  
 RX MEDLINE; 93224141.  
 RA NONAKA M., MATSUDA Y., SHIROISHI T., MORIWAKI K., NONAKA M.,  
 RA NATSUUME-SAKAI S.;  
 RT "Molecular cloning of the b subunit of mouse coagulation factor XIII  
 RT and assignment of the gene to chromosome 1: close evolutionary  
 RT relationship to complement factor H.";  
 RL GENOMICS 15:535-542(1993).  
 CC -!- FUNCTION: THE B CHAIN OF FACTOR XIII IS NOT CATALYTICALLY ACTIVE,  
 CC BUT IS THOUGHT TO STABILIZE THE A SUBUNITS AND REGULATE THE RATE  
 CC OF TRANSGLUTAMINASE FORMATION BY THROMBIN.  
 CC -!- SUBUNIT: TETRAMER OF TWO A CHAINS AND TWO B CHAINS.  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN LIVER AND KIDNEY.  
 CC -!- SIMILARITY: CONTAINS 10 SUSHI (SCR) REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D10071; G303652; -  
 DR PIR; A46013; A46013.  
 DR MGD; MGI:88379; CF13B.  
 DR PFAM; PF00084; sushi; 8.  
 DR HSP; P10998; 1VVC.  
 KW TRANSFERASE; PLASMA; BLOOD COAGULATION; REPEAT; GLYCOPROTEIN; SIGNAL;  
 KW SUSHI.  
 FT SIGNAL. 1 20 BY SIMILARITY.  
 FT CHAIN 21 668 FACTOR XIII, B CHAIN.  
 FT DOMAIN 24 647 10 X SUSHI (SCR) REPEATS.  
 FT REPEAT 24 88 SUSHI 1.  
 FT REPEAT 90 147 SUSHI 2.  
 FT REPEAT 152 209 SUSHI 3.  
 FT REPEAT 212 268 SUSHI 4.  
 FT REPEAT 273 328 SUSHI 5.

FT REPEAT 335 390 SUSHI 6.  
 FT REPEAT 395 451 SUSHI 7.  
 FT REPEAT 453 516 SUSHI 8.  
 FT REPEAT 523 579 SUSHI 9.  
 FT REPEAT 581 647 SUSHI 10.  
 FT DISULFID 25 76 BY SIMILARITY.  
 FT DISULFID 59 87 BY SIMILARITY.  
 FT DISULFID 91 135 BY SIMILARITY.  
 FT DISULFID 118 146 BY SIMILARITY.  
 FT DISULFID 153 197 BY SIMILARITY.  
 FT DISULFID 180 208 BY SIMILARITY.  
 FT DISULFID 213 255 BY SIMILARITY.  
 FT DISULFID 241 267 BY SIMILARITY.  
 FT DISULFID 274 316 BY SIMILARITY.  
 FT DISULFID 302 327 BY SIMILARITY.  
 FT DISULFID 336 378 BY SIMILARITY.  
 FT DISULFID 364 389 BY SIMILARITY.  
 FT DISULFID 396 439 BY SIMILARITY.  
 FT DISULFID 425 450 BY SIMILARITY.  
 FT DISULFID 454 505 BY SIMILARITY.  
 FT DISULFID 486 515 BY SIMILARITY.  
 FT DISULFID 524 567 BY SIMILARITY.  
 FT DISULFID 553 578 BY SIMILARITY.  
 FT DISULFID 582 636 BY SIMILARITY.  
 FT DISULFID 616 646 BY SIMILARITY.  
 FT CARBOHYD 162 162 POTENTIAL.  
 FT CARBOHYD 545 545  
 SQ SEQUENCE 668 AA; 76078 MW; 45234BA0 CRC32;  
 Query Match 70.3%; Score 45; DB 1; Length 668;  
 Best Local Similarity 77.8%; Pred. No. 1.6;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 RLLWKYKGF 10  
 :|||||:  
 Db 467 QLKWKYEGK 475  
 RESULT 4  
 F13B\_HUMAN STANDARD; PRT; 661 AA.  
 ID F13B\_HUMAN  
 AC P05160;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE COAGULATION FACTOR XIII B CHAIN PRECURSOR (EC 2.3.2.13) (PROTEIN-  
 DE GLUTAMINE GAMMA-GLUTAMYLTRANSFERASE B CHAIN) (TRANSGLUTAMINASE B  
 DE CHAIN).  
 GN F13B.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91105054.  
 RA BOTTENUS R.E., ICHINOSE A., DAVIE E.W.;  
 RT "Nucleotide sequence of the gene for the b subunit of human factor  
 RT XIII.";  
 RL BIOCHEMISTRY 29:11195-11209(1990).  
 RN [2]  
 RP SEQUENCE OF 2-661 FROM N.A.  
 RX MEDLINE; 87026535.  
 RA ICHINOSE A., McMULLEN B.A., FUJIKAWA K., DAVIE E.W.;  
 RT "Amino acid sequence of the b subunit of human factor XIII, a protein  
 RT composed of ten repetitive segments.";  
 RL BIOCHEMISTRY 25:4633-4638(1986).  
 RN [3]  
 RP REVISIONS.  
 RA ICHINOSE A.;  
 RL SUBMITTED (FEB-1987) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [4]  
 RP SEQUENCE OF 1-20 FROM N.A.

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:08 ; Search time 71.09 Seconds  
(without alignments)  
4.374 Million cell updates/sec

Title: US-09-124-280A-28

Perfect score: 64

Sequence: 1 KRLWKYKGF 11

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	64	100.0	102	1	ALPS_TACTR	P07087 tachypleus
2	61	95.3	101	1	ALPS_LIMPO	P07086 limulus pol
3	45	70.3	668	1	F13B_MOUSE	Q07968 mus musculus
4	42	65.6	661	1	F13B_HUMAN	P05160 homo sapien
5	41	64.1	563	1	HNML_YEAST	P19807 saccharomyc
6	40	62.5	2222	1	DPOE_YEAST	P21951 saccharomyc
7	40	62.5	277	1	HXDB_NOTVI	P31263 notophthalm
8	39	60.9	2029	1	LARDROME	P16621 drosophila
9	39	60.9	106	1	YQR5_YEAST	P32339 saccharomyc
10	38	59.4	323	1	CD47_HUMAN	Q08722 homo sapien
11	38	59.4	215	1	RL10_EUGGR	Q39724 euglena gra
12	38	59.4	527	1	UDAL_RAT	P36510 rattus norv
13	37	57.8	280	1	HXDB_CHICK	P24342 gallus gall
14	37	57.8	347	1	NU2M_DASNO	O21326 dasyneus nov
15	37	57.8	273	1	PSAB_YERPE	P04157 rattus norv
16	36	56.2	1255	1	CD45_RAT	P26619 xenopus lae
17	36	56.2	1087	1	PGDS_XENLA	P28973 equine herp
18	36	56.2	766	1	PRTP_HSVB	P46756 acanthamoeb
19	36	56.2	337	1	RT07_ACACA	P23813 mus musculus
20	35	54.7	323	1	HXDB_MOUSE	P35845 saccharomyc
21	35	54.7	859	1	OSHL_YEAST	O08357 sulfolobus
22	35	54.7	355	1	PIRC_SULSO	P50621 bacillus su
23	35	54.7	329	1	RIR2_BACSU	P13807 homo sapien
24	35	54.7	737	1	UGS1_HUMAN	P54859 mus musculus
25	35	54.7	737	1	UGS3_MOUSE	P27974 simian immu
26	35	54.7	232	1	VIF_SIVAI	Q57889 methanococc
27	35	54.7	161	1	Y447_METJA	P41647 pinus thunb
28	35	54.7	1756	1	YCF1_PINTH	P53266 saccharomyc
29	35	54.7	389	1	YG2X_YEAST	P47134 saccharomyc
30	35	54.7	954	1	XJ59_YEAST	O43143 homo sapien
31	34.5	53.9	813	1	DD13_HUMAN	O35286 mus musculus
32	34.5	53.9	758	1	DD13_MOUSE	Q35286 mus musculus
33	34	53.1	401	1	ACKA_HAEIN	P44406 haemophilus
34	34	53.1	232	1	CSCI_ECOLI	P53518 escherichia
35	34	53.1	232	1	CSCI_ECOLI	P53519 escherichia
36	34	53.1	2257	1	DPOE_HUMAN	Q07864 homo sapien
37	34	53.1	220	1	HB7A_XENLA	Q91771 xenopus lae
38	34	53.1	220	1	HB7B_XENLA	P04476 xenopus lae
39	34	53.1	113	1	HEMT_PHAGO	P02244 phascoplopsi
40	34	53.1	113	1	HEMT_STPCU	P22766 siphonoma
41	34	53.1	113	1	HEMT_THREZO	P02245 themiste zo
42	34	53.1	378	1	HMAN_DROME	P02833 drosophila
43	34	53.1	66	1	HXA6_HUMAN	P31267 homo sapien

44 34 53.1 97 1 HXA6\_MOUSE P09092 mus musculus  
45 34 53.1 228 1 HXB6\_BRARE P15861 brachydanio

ALIGNMENTS

```
RESULT 1
ALPS_TACTR
ID ALPS_TACTR STANDARD; PRT; 102 AA.
AC P07087;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE ANTI-LIPOPOLYSACCHARIDE FACTOR (ANTI-LPS).
OS TACHYPLEUS TRIDENTATUS (JAPANESE HORSESHOE CRAB).
OC EUKARYOTA; METAZOA; ARTHROPODA; CHELICERATA; XIPHOSURA;
LN LIMULIDAE; TACHYPLEUS.
RN [1]
RP SEQUENCE.
RX MEDLINE: 86224003.
RA AKETAGANA J., MIYATA T., OHTSUBO S., NAKAMURA T., MORITA T.,
RA HAYASHIDA H., IWANAGA S., TAKAO T., SHIMONISHI Y.;
RT "Primary structure of limulus anticoagulant anti-lipopolysaccharide
RT factor.";
RL J. BIOL. CHEM. 261:7357-7365(1986).
CC -I- FUNCTION: BINDS TIGHTLY TO LPS AND THUS SPECIFICALLY INHIBITS THE
CC LPS-MEDIATED ACTIVATION OF THE HEMOLYMPH COAGULATION. IT HAS A
CC STRONG ANTIBACTERIAL EFFECT ESPECIALLY ON THE GROWTH OF GRAM-
CC NEGATIVE BACTERIA.
CC PIR; A23931; A23931.
DR ANTI-BIOTIC.
KW DISULFID 32 53
FT VARIANT 36 36 V -> I.
FT VARIANT 102 102 E -> Q.
SQ SEQUENCE 102 AA; 11596 MW; 5823800C CRC32;

Query Match 100.0%; Score 64; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLWKYKGF 11
DB 41 KRLWKYKGF 51
|||||
|||||

RESULT 2
ALPS_LIMPO
ID ALPS_LIMPO STANDARD; PRT; 101 AA.
AC P07086;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE ANTI-LIPOPOLYSACCHARIDE FACTOR (ANTI-LPS).
OS LIMULUS POLYSPHEMUS (ATLANTIC HORSESHOE CRAB).
OC EUKARYOTA; METAZOA; ARTHROPODA; CHELICERATA; XIPHOSURA;
LN LIMULIDAE; LIMULUS.
RN [1]
RP SEQUENCE.
RX MEDLINE: 88032893.
RA MUTA T., MIYATA T., TOKUNAGA F., NAKAMURA T., IWANAGA S.;
RT "Primary structure of anti-lipopolysaccharide factor from American
RT horseshoe crab, Limulus polyphemus.";
RL J. BIOCHEM. 101:1321-1330(1987).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE: 94074533.
RA HOESS A., WATSON S., SIBER G.R., LIDDINGTON R.;
RT "Crystal structure of an endotoxin-neutralizing protein from the
RT horseshoe crab, Limulus anti-LPS factor, at 1.5-A resolution.";
RL EMBO J. 12:3351-3356(1993).
CC -I- FUNCTION: BINDS TIGHTLY TO LPS AND THUS SPECIFICALLY INHIBITS THE
```

**THIS PAGE BLANK (USPTO)**

Best Local Similarity 55.6%; Pred. No. 13;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LKWKYKGF 11  
| | | | |  
Db 96 LSWKFRGEF 104

## RESULT 14

D48997  
tumor surface antigen OA3-293 - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997  
C:Accession: D48997  
R:Campbell, I.G.; Freemont, P.S.; Foulkes, W.; Trowsdale, J.  
A:Title: An ovarian tumor marker with homology to vaccinia virus contains an IgV-like re  
Cancer Res. 52, 5416-5420, 1992  
A:Reference number: A48997; MUID:93007897  
A:Accession: D48997  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-293 <CAM>  
A:Experimental source: ovarian carcinoma line PE/01  
A>Note: sequence extracted from NCBI backbone (NCBIN:114647, NCBIP:114652)  
C:Keywords: surface antigen

Query Match 59.4%; Score 38; DB 2; Length 293;  
Best Local Similarity 62.5%; Pred. No. 23;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LKWKYKGF 10  
| | | | |  
Db 56 VKWKFGR 63

## RESULT 15

C48997  
tumor surface antigen OA3-305 - human  
N:Alternate names: integrin-associated protein  
C:Species: Homo sapiens (man)  
C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Sep-1997  
C:Accession: C48997; S36644  
R:Campbell, I.G.; Freemont, P.S.; Foulkes, W.; Trowsdale, J.  
A:Title: An ovarian tumor marker with homology to vaccinia virus contains an IgV-like re  
Cancer Res. 52, 5416-5420, 1992  
A:Reference number: A48997; MUID:93007897  
A:Accession: C48997  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-305 <CAM>  
A:Experimental source: ovarian carcinoma line PE/01  
A>Note: sequence extracted from NCBI backbone (NCBIN:114645, NCBIP:114651)  
R:Lindberg, P.  
submitted to the EMBL Data Library, August 1993  
A:Reference number: S36644  
A:Accession: S36644  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-305 <LIN>  
A:Cross-references: EMBL:225521; NID:g396704; PID:g396705  
C:Keywords: surface antigen

Query Match 59.4%; Score 38; DB 2; Length 305;  
Best Local Similarity 62.5%; Pred. No. 24;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LKWKYKGF 10  
| | | | |  
Db 56 VKWKFGR 63

Search completed: September 7, 1999, 23:22:43  
Job time: 950 sec

hypothetical protein BB0170 - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
 C:Accession: B70121  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, S.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, B.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: B70121  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-682 <LE>  
 A:Cross-references: GB:AE001128; GB:AE000783; NID:g2688057; PID:g2688070; TIGR:BB0170  
 A:Experimental source: strain B31

Query Match 60.9%; Score 39; DB 2; Length 682;  
 Best Local Similarity 54.5%; Pred. No. 36;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
 :||:| |:  
 DB 369 KKLKYEYSGEF 379

RESULT 10  
 S48556  
 Probable membrane protein YLR205c - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein L8167.18  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 13-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 17-Mar-1999  
 C:Accession: S48556; S25342  
 R:Pauley, A.  
 submitted to the EMBL Data Library, September 1994  
 A:Description: The sequence of S. cerevisiae cosmid 8167.  
 A:Reference number: S48545  
 A:Accession: S48556  
 A:Molecule type: DNA  
 A:Residues: 1-273 <PAU>  
 A:Cross-references: EMBL:U14913; NID:g544497; PID:g544514; MIPS:YLR205c  
 R:Simon, M.; della Seta, F.; Sor, F.; Faye, G. Yeast 8, 559-567, 1992  
 A:Title: Analysis of the MSS51 region on chromosome XII of Saccharomyces cerevisiae.  
 A:Reference number: S25342; MUID:92397593  
 A:Accession: S25342  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 78-83, 'H', '85-181, 'RN' <SIM>  
 C:Genetics:  
 A:Map position: 12R  
 C:Keywords: transmembrane protein  
 F:105-121/Domain: transmembrane #status predicted <TM1>  
 F:249-265/Domain: transmembrane #status predicted <TM2>

Query Match 60.9%; Score 39; DB 2; Length 273;  
 Best Local Similarity 60.0%; Pred. No. 14;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RLKWKYKGF 11  
 |||||:  
 DB 167 RLKWKYKKNY 176

RESULT 11  
 S39068  
 homeotic protein cnorx5 - Hydra magnipapillata (fragment)  
 C:Species: Hydra magnipapillata  
 C:Date: 07-Oct-1994 #sequence\_revision 30-Jan-1998 #text\_change 30-Jan-1998  
 C:Accession: S39068

R:Naito, M.; Ishiguro, H.; Fujisawa, T.; Kurosawa, Y. FEBS Lett. 333, 271-274, 1993  
 A:Title: Presence of eight distinct homeobox-containing genes in cnidarians.  
 A:Reference number: S39066; MUID:94039774  
 A:Accession: S39068  
 A:Molecule type: nucleic acid  
 A:Residues: 1-59 <NAI>  
 A:Cross-references: EMBL:222640  
 C:Genetics:  
 A:Gene: cnorx5  
 C:Superfamily: unassigned homeobox proteins; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
 F:1-43/Domain: homeobox homology (fragment) <HOX>

Query Match 59.4%; Score 38; DB 2; Length 59;  
 Best Local Similarity 50.0%; Pred. No. 4.4;  
 Matches 7; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

QY 1 KRLKWK----YKKG 10  
 :|:|:| |:  
 DB 37 RRKWKRVKGYRKG 50

RESULT 12  
 S15089  
 glucuronosyltransferase (EC 2.4.1.17) - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 13-Sep-1998  
 C:Accession: S15089  
 R:Lazard, D.; Zupko, K.; Poria, Y.; Nef, P.; Lazarovits, J.; Horn, S.; Khen, M.; Lanc Nature 349, 790-793, 1991  
 A:Title: Odorant signal termination by olfactory UDP glucuronosyl transferase.  
 A:Reference number: S15089; MUID:91156050  
 A:Accession: S15089  
 A:Molecule type: mRNA  
 A:Residues: 1-527 <LAZ>  
 C:Superfamily: glucuronosyltransferase  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 59.4%; Score 38; DB 2; Length 527;  
 Best Local Similarity 50.0%; Pred. No. 41;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYKKG 10  
 :|:|:| |:  
 DB 331 QKVLWRYKKG 340

RESULT 13  
 E69119  
 conserved hypothetical protein MTH1892 - Methanobacterium thermoautotrophicum (strain C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
 C:Accession: E69119  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafore, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jilwan, K.I.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu  
 A:Reference number: A69000; MUID:98037514  
 A:Accession: E69119  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-168 <MTH>  
 A:Cross-references: GB:AE000940; GB:AE000666; NID:g2623011; PID:g2623024  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH1892

Query Match 59.4%; Score 38; DB 2; Length 168;

C:Species: Saccharomyces cerevisiae  
 C:Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 17-Mar-1999  
 C:Accession: A36028; B36028; S60919; S63235; S65121  
 R:Morrison, A.; Araki, H.; Clark, A.B.; Hamatake, R.K.; Sugino, A.  
 Cell 62, 1143-1151, 1990  
 A:Title: A third essential DNA polymerase in Saccharomyces cerevisiae.  
 A:Reference number: A36028; MUID:90381771  
 A:Accession: A36028  
 A:Molecule type: DNA  
 A:Residues: 1-2222 <MOR>  
 A:Cross-references: GB:M60416; GB:M36724; NID:gl711408; PID:gl711409  
 A:Accession: B36028  
 A:Molecule type: protein  
 A:Residues: 1214-1216, X', 1218-1221 <MO2>  
 R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.  
 submitted to The EMBL Data Library, October 1995  
 A:Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV fr  
 A:Reference number: S60909  
 A:Accession: S60919  
 A:Molecule type: DNA  
 A:Residues: 1-2221 <SEN>  
 A:Cross-references: EMBL:X92494; NID:gl045236; PID:gl045247  
 R:Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.  
 submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S63235  
 A:Accession: S63235  
 A:Molecule type: DNA  
 A:Residues: 1-2222 <SEW>  
 A:Cross-references: EMBL:271538; NID:gl1302316; PID:e239648; PID:gl1302317; MIPS:YNL262w  
 A:Experimental source: strain S288C  
 R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.  
 Yeast 12, 505-514, 1996  
 A:Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa  
 A:Reference number: S65111; MUID:96310631  
 A:Accession: S65121  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2221 <SEF>  
 A:Cross-references: EMBL:X92494; NID:gl045236; PID:gl045247  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
 C:Genetics:  
 A:Gene: SGD:POL2; DUN2  
 A:Cross-references: SGD:S0005206; MIPS:YNL262w  
 A:Map position: 14L  
 C:Superfamily: DNA-directed DNA polymerase II  
 C:Keywords: DNA binding; nucleotidyltransferase; nucleus; zinc finger

Query Match 62.5%; Score 40; DB 2; Length 2222;  
 Best Local Similarity 45.5%; Pred. No. 82;  
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
 Db 679 RKLKAWRGGEF 689

RESULT 7  
 A43783  
 homeotic protein NvHox2 - eastern newt  
 C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)  
 C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 17-Oct-1997  
 C:Accession: A43783  
 R:Brown, R.; Brookes, J.P.  
 Development 111, 489-496, 1991  
 A:Title: Identification and expression of a regeneration-specific homeobox gene in the n  
 A:Reference number: A43783; MUID:91372146  
 A:Accession: A43783  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-277 <BRO>  
 A:Cross-references: GB:X57674  
 C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
 F:206-262/Domain: homeobox homology <HOX>

Query Match 62.5%; Score 40; DB 2; Length 277;  
 Best Local Similarity 45.5%; Pred. No. 9.8;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
 Db 68 RRSKQYRGSY 78

RESULT 8

TOFFLK

protein-tyrosine-phosphatase (EC 3.1.3.48) DLAR precursor - fruit fly (Drosophila mel  
 N:Alternate names: leukocyte antigen-related protein  
 C:Species: Drosophila melanogaster  
 C:Date: 14-Dec-1990 #sequence\_revision 02-May-1994 #text\_change 12-Feb-1999  
 C:Accession: A36182  
 R:Streuli, M.; Krueger, N.X.; Tsai, A.Y.M.; Saito, H.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 8698-8702, 1989  
 A:Title: A family of receptor-linked protein tyrosine phosphatases in humans and Dros

A:Reference number: A36182; MUID:90046860  
 A:Accession: A36182  
 A:Molecule type: mRNA  
 A:Residues: 1-2029 <STR>  
 A:Cross-references: GB:M27700; NID:gl57811; PID:gl57812  
 C:Genetics:  
 A:Gene: FlyBase:Lar

A:Cross-references: FlyBase:FBgn0000464

C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homolog

C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembra

F:1-32/Domain: signal sequence #status predicted <Sig>  
 F:33-2029/Product: leukocyte antigen-related protein #status predicted <MAT>  
 F:33-1377/Domain: extracellular #status predicted <EXT>

F:50-113/Domain: immunoglobulin homology <IMM1>

F:154-211/Domain: immunoglobulin homology <IMM2>

F:249-303/Domain: immunoglobulin homology <IMM3>

F:321-401/Domain: fibronectin type III repeat homology <FN3A>

F:416-502/Domain: fibronectin type III repeat homology <FN3B>

F:514-599/Domain: fibronectin type III repeat homology <FN3C>

F:610-699/Domain: fibronectin type III repeat homology <FN3D>

F:708-802/Domain: fibronectin type III repeat homology <FN3E>

F:811-896/Domain: fibronectin type III repeat homology <FN3F>

F:909-993/Domain: fibronectin type III repeat homology <FN3G>

F:1006-1091/Domain: fibronectin type III repeat homology <FN3H>

F:1101-1198/Domain: fibronectin type III repeat homology <FN3I>

F:1378-1402/Domain: transmembrane #status predicted <TM>

F:1403-2029/Domain: intracellular #status predicted <INT>

F:1417-2029/Domain: leukocyte common antigen cytosolic domain homology <LAC>

F:1497-1718/Domain: protein-tyrosine-phosphatase homology <PTP1>

F:1786-2009/Domain: protein-tyrosine-phosphatase homology <PTP2>

F:57-111,161-209,256-301/Disulfide bonds: #status predicted

F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:1670/Active site: Cys (phosphocysteine intermediate) #status predicted

F:1676/Binding site: substrate phosphate (Arg) #status predicted

F:1961/Active site: Cys (phosphocysteine intermediate) #status predicted

F:1967/Binding site: substrate phosphate (Arg) #status predicted

Query Match 60.9%; Score 39; DB 1; Length 2029;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RLKWKYKG 9

Db 339 RLKWKYKG 346

RESULT 9

B70121



A:Molecule type: nucleic acid  
A:Residues: 1-668 <NON>  
A:Cross-references: GB:D10071; NID:g303651; PID:d1001429; PID:g303652  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCB129405, NCBIP:129407)  
C:Superfamily: coagulation factor XIII chain b; complement factor H repeat homology  
F:25-87/Domain: complement factor H repeat homology <FH01>  
F:91-146/Domain: complement factor H repeat homology <FH02>  
F:153-208/Domain: complement factor H repeat homology <FH03>  
F:213-267/Domain: complement factor H repeat homology <FH04>  
F:274-327/Domain: complement factor H repeat homology <FH05>  
F:336-389/Domain: complement factor H repeat homology <FH06>  
F:396-450/Domain: complement factor H repeat homology <FH07>  
F:454-515/Domain: complement factor H repeat homology <FH08>  
F:582-646/Domain: complement factor H repeat homology <FH10>

Query Match 70.3%; Score 45; DB 2; Length 668;  
Best Local Similarity 77.8%; Pred. No. 3.4;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 RLKWKYK 10  
:|||||:  
Db 467 QLKWKYK 475

RESULT 4  
KFHUI3  
coagulation factor XIII chain b precursor - human  
N:Alternate names: fibrin-stabilizing factor; plasma transglutaminase b chain; protein-g  
C:Species: Homo sapiens (man)  
C>Date: 19-Apr-1991 #sequence\_revision 26-May-1994 #text\_change 20-Mar-1998  
C:Accession: A36397; A23830; S09980  
R:Botteon, R.E.; Ichinose, A.; Davie, E.W.  
Biochemistry 29, 11195-11209, 1990  
A:Title: Nucleotide sequence of the gene for the b subunit of human factor XIII.  
A:Reference number: A36397; MUID:91105054  
A:Accession: A36397  
A:Molecule type: DNA  
A:Residues: 1-661 <BOT>  
A:Cross-references: GB:J05294  
R:Ichinose, A.; McMullen, B.A.; Fujikawa, K.; Davie, E.W.  
Biochemistry 25, 4633-4638, 1986  
A:Title: Title: Amino acid sequence of the b subunit of human factor XIII, a protein com  
A:Reference number: A23830; MUID:87026535  
A:Accession: A23830  
A:Molecule type: mRNA  
A:Residues: 2-661 <ICH>  
A:Cross-references: GB:M14057; NID:g182838; PID:g182839  
A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
R:Grundmann, U.; Nerlich, C.; Rein, T.; Zettlmeissl, G.  
Nucleic Acids Res. 18, 2817-2818, 1990  
A:Title: Complete cDNA sequence encoding the B subunit of human factor XIII.  
A:Reference number: S09980; MUID:90251467  
A:Accession: S09980  
A:Molecule type: mRNA  
A:Residues: 1-661 <GRU>  
A:Cross-references: EMBL:X51823  
C:Comment: This protein is noncatalytic.  
C:Genetics:  
A:Gene: GDB:F13B  
A:Cross-references: GDB:119893; OMIM:134580  
A:Map position: 1q31-lq32.1  
A:Introns: 2/1; 69/1; 131/1; 190/1; 249/1; 309/1; 371/1; 432/1; 499/1; 560/1; 631/2  
C:Complex: heterotetramer in plasma with coagulation factor XIII chain a, the plasma tra  
C:Superfamily: coagulation factor XIII chain b; complement factor H repeat homology  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-661/Product: coagulation factor XIII chain b #status predicted <MAT>  
F:25-87/Domain: complement factor H repeat homology <FH01>  
F:91-146/Domain: complement factor H repeat homology <FH02>  
F:153-208/Domain: complement factor H repeat homology <FH03>  
F:213-267/Domain: complement factor H repeat homology <FH04>

F:274-327/Domain: complement factor H repeat homology <FH05>  
F:336-389/Domain: complement factor H repeat homology <FH06>  
F:396-450/Domain: complement factor H repeat homology <FH07>  
F:454-515/Domain: complement factor H repeat homology <FH08>  
F:524-578/Domain: complement factor H repeat homology <FH09>  
F:582-646/Domain: complement factor H repeat homology <FH10>  
F:617-619/Region: cell attachment (R-G-D) motif  
F:25-76,59-87,91-135,118-146,153-197,180-208,213-255,241-267,274-316,302-327,336-378,  
F:162,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 65.68; Score 42; DB 1; Length 661;  
Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LKWKYK 10  
:|||||:  
Db 468 MKWKYK 475

RESULT 5  
S11175  
choline transport protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein G313; protein YGL077c  
C:Species: Saccharomyces cerevisiae  
C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 05-Dec-1997  
C:Accession: S11175; S64084  
R:Nikawa, J.I.; Hosaka, K.; Tsukagoshi, Y.; Yamashita, S.  
J. Biol. Chem. 265, 15996-16003, 1990  
A:Title: Primary structure of the yeast choline transport gene and regulation of its  
A:Reference number: S11175; MUID:90368823  
A:Accession: S11175  
A:Molecule type: DNA  
A:Residues: 1-563 <NIK>  
A:Cross-references: EMBL:J05603; NID:g171329; PID:g171330  
R:Reger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64071  
A:Accession: S64084  
A:Molecule type: DNA  
A:Residues: 1-563 <RIE>  
A:Cross-references: EMBL:Z72599; NID:g1322592; PID:e243965; PID:g1322593; MIPS:YGL077  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:HNM1; CTR1  
A:Cross-references: MIPS:YGL077c; SGD:S0003045  
A:Map position: 7L  
C:Superfamily: choline transport protein  
C:Keywords: transmembrane protein  
F:91-107/Domain: transmembrane #status predicted <TM1>  
F:185-201/Domain: transmembrane #status predicted <TM2>  
F:214-230/Domain: transmembrane #status predicted <TM3>  
F:257-273/Domain: transmembrane #status predicted <TM4>  
F:298-314/Domain: transmembrane #status predicted <TM5>  
F:345-361/Domain: transmembrane #status predicted <TM6>  
F:401-417/Domain: transmembrane #status predicted <TM7>  
F:428-444/Domain: transmembrane #status predicted <TM8>  
F:466-482/Domain: transmembrane #status predicted <TM9>

Query Match 64.1%; Score 41; DB 2; Length 563;  
Best Local Similarity 87.5%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 LKWKYK 10  
:|||||:  
Db 508 LYWKYK 515

RESULT 6  
A36028  
DNA-directed DNA polymerase (EC 2.7.7.7) II catalytic chain - yeast (Saccharomyces ce  
N:Alternate names: DNA-directed DNA polymerase II chain A; protein NO825; protein YNL

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:42 ; Search time 105.14 Seconds  
(without alignments)  
4.192 Million cell updates/sec

Title: US-09-124-280A-28  
Perfect score: 64  
Sequence: 1 KRLKWKYKGF 11

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	102	2 A23931	anti-lipoplysacch
2	61	95.3	101	2 A27819	anti-lipoplysacch
3	45	70.3	668	2 A46013	coagulation factor
4	42	65.6	661	1 KFHU13	coagulation factor
5	41	64.1	563	2 S11175	choline transport
6	40	62.5	2222	2 A36038	DNA-directed DNA p
7	40	62.5	277	2 A43783	homeotic protein N
8	39	60.9	2029	1 TDFELK	protein-tyrosine-p
9	39	60.9	682	2 B70121	hypothetical prote
10	39	60.9	273	2 S48536	probable membrane
11	38	59.4	59	2 S39068	homeotic protein c
12	38	59.4	527	2 S15089	glucuronosyltransf
13	38	59.4	168	2 E69119	conserved hypothet
14	38	59.4	293	2 D48997	tumor surface anti
15	38	59.4	305	2 C48997	tumor surface anti
16	38	59.4	312	2 B48997	tumor surface anti
17	38	59.4	323	2 A48997	tumor surface anti
18	37	57.8	273	2 S32927	psab protein precu
19	36	56.2	1246	1 TDRTUT	leukocyte common a
20	36	56.2	766	1 WZBEC5	infected cell prot
21	36	56.2	445	1 P71033	hypothetical prote
22	36	56.2	1087	2 I51532	platelet-derived g
23	36	56.2	287	2 A45854	leucocyte common a
24	36	56.2	295	2 H71085	probable MRP prote
25	36	56.2	398	2 E64913	hypothetical prote
26	36	56.2	253	2 D71682	cytochrome c1, hem
27	36	56.2	480	2 F69774	transposon-related
28	36	56.2	957	2 H69141	hypothetical prote
29	36	56.2	337	2 S53846	ribosomal protein
30	36	56.2	438	2 D71175	hypothetical prote
31	36	56.2	758	2 D71072	hypothetical prote
32	35	54.7	323	2 S16318	homeotic protein H
33	35	54.7	329	2 C59667	ribonucleoside-dip
34	35	54.7	337	2 S74850	hypothetical prote
35	35	54.7	168	2 S77282	hypothetical prote
36	35	54.7	226	2 S32215	hypothetical prote
37	35	54.7	195	2 B69937	ribonucleoside-dip
38	35	54.7	161	2 G54355	hypothetical prote
39	35	54.7	954	2 S57108	hypothetical prote

ALIGNMENTS

```
RESULT 1
A23931
anti-lipoplysaccharide factor - Atlantic horseshoe crab
C:Species: Limulus polyphemus (Atlantic horseshoe crab)
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1993
C:Accession: A23931
R:Aketagawa, J.; Miyata, T.; Ohtsubo, S.; Nakamura, T.; Morita, T.; Hayashida, H.; Mi
J. Biol. Chem. 261, 7357-7365, 1986
A:Title: Primary structure of Limulus anticoagulant anti-lipoplysaccharide factor.
A:Reference number: A23931; MUID:86224003
A:Accession: A23931
A:Molecule type: protein
A:Residues: 1-102 <AKE>

Query Match 100.0%; Score 64; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11
|||||
Db 41 KRLKWKYKGF 51

RESULT 2
A27819
anti-lipoplysaccharide factor - Atlantic horseshoe crab
C:Species: Limulus polyphemus (Atlantic horseshoe crab)
C>Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 18-Jun-1993
C:Accession: A27819
R:Muta, T.; Miyata, T.; Tokunaga, F.; Nakamura, T.; Iwanaga, S.
J. Biochem. 101, 1321-1330, 1987
A:Title: Primary structure of anti-lipoplysaccharide factor from American horseshoe
A:Reference number: A27819; MUID:88032893
A:Accession: A27819
A:Molecule type: protein
A:Residues: 1-101 <MUT>
A:Note: 13-Lys was also found
F:31-52/Disulfide bonds: #status experimental

Query Match 95.3%; Score 61; DB 2; Length 101;
Best Local Similarity 90.9%; Pred. No. 0.00095;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11
:|||||
Db 40 RRLKWKYKGF 50

RESULT 3
A46013
coagulation factor XIII subunit b - mouse
C:Species: Mus musculus (house mouse)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 16-Dec-1998
C:Accession: A46013
R:Nonaka, M.; Matsuda, Y.; Shiroishi, T.; Moriwaki, K.; Nonaka, M.; Natsume-Sakai, S
Genomics 15, 535-542, 1993
A:Title: Molecular cloning of the b subunit of mouse coagulation factor XIII and assi
A:Reference number: A46013; MUID:93224141
A:Accession: A46013
A>Status: preliminary
```

```
40 35 54.7 389 2 S64420 mitochondrial inne
41 35 54.7 1246 2 S60954 probable membrane
42 35 54.7 1190 2 S47536 SWH1 protein - yea
43 35 54.7 737 2 A32156 UDPglucose--glycog
44 35 54.7 735 2 A33369 UDPglucose--glycog
45 35 54.7 558 2 F71175 hypothetical prote
```

**THIS PAGE BLANK (USPTO)**

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/366,953A  
FILING DATE: 30-DEC-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Cook Ph.D., Robert R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-324  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-366-953A-25

Query Match 92.2%; Score 59; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00099;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLKWKYKGF 11  
Db 1 RLKWKYKGF 10

Search completed: September 7, 1999, 23:07:47  
Job time: 1745 sec

REFERENCE/DOCKET NUMBER: 1413.0010005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
TELEX: 248636 SSK  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-704-872-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
Db 40 RRLKWKYKGF 50

RESULT 13  
US-08-704-872-3

Sequence 3, Application US/08704872  
Patent No. 5747455  
GENERAL INFORMATION:  
APPLICANT: Wainwright, No. 5747455man R.  
TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
TITLE OF INVENTION: Protein and Uses Thereof  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: United States of America  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704.872  
FILING DATE: 30-AUG-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/264,244  
FILING DATE: 22-JUNE-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1413.0010005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
TELEX: 248636 SSK  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-704-872-3

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
Db 40 RRLKWKYKGF 50

Db 44 RRLKWKYKGF 54

RESULT 14

US-08-577-464-2  
Sequence 2, Application US/08577464  
Patent No. 5780429  
GENERAL INFORMATION:  
APPLICANT: WAINWRIGHT, NO. 5780429man R.  
TITLE OF INVENTION: ANTI-LPS FACTOR FROM HORSESHOE CRABS AND  
TITLE OF INVENTION: METHODS OF USE  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr LLP  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/577.464  
FILING DATE: 22-DEC-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: BAKER, HOLLIE L.  
REGISTRATION NUMBER: 31,321  
REFERENCE/DOCKET NUMBER: 58378.650  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-577-464-2

Query Match 95.3%; Score 61; DB 2; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
Db 40 RRLKWKYKGF 50

RESULT 15

US-08-366-953A-25  
Sequence 25, Application US/08366953A  
Patent No. 5766593  
GENERAL INFORMATION:  
APPLICANT: Lichenstein, Henri S.  
APPLICANT: Wright, Samuel D.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 PEPTIDES  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1720  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

QY 1 KRLKWKYKGF 11  
:|||||  
Db 44 RRLKWKYKGF 54

RESULT 10  
US-08-476-940-1  
; Sequence 1, Application US/08476940  
; Patent No. 5627266  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5627266man R.  
; APPLICANT: No. 5627266itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses Thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,940  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/264,244  
; FILING DATE: 22-JUNE-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1413.0010005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; TELEX: 248636 SSK  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-476-940-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
Db 40 RRLKWKYKGF 50

RESULT 11  
US-08-476-940-3  
; Sequence 3, Application US/08476940  
; Patent No. 5627266  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5627266man R.  
; APPLICANT: No. 5627266itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses Thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.

; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,940  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/264,244  
; FILING DATE: 22-JUNE-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1413.0010005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; TELEX: 248636 SSK  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-476-940-3

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
Db 44 RRLKWKYKGF 54

RESULT 12  
US-08-704-872-1  
; Sequence 1, Application US/08704872  
; Patent No. 5747455  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5747455man R.  
; APPLICANT: No. 5747455itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses Thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,872  
; FILING DATE: 30-AUG-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/264,244  
; FILING DATE: 22-JUNE-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893

ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1413.0010005  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
TELEX: 248636 SSK  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-478-689-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
Db 40 RRLKWKYKGF 50

RESULT 9  
US-08-478-689-3  
Sequence 3, Application US/08478689  
Patent No. 5614369  
GENERAL INFORMATION:  
APPLICANT: Wainwright, No. 5614369man R.  
APPLICANT: No. 5614369itsky, Thomas J.  
TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
TITLE OF INVENTION: Protein and Uses Thereof  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: United States of America  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/264,244  
FILING DATE: 22-JUNE-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1413.0010005  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
TELEX: 248636 SSK  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-264-244-3

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
Db 44 RRLKWKYKGF 54

RESULT 8  
US-08-478-689-1  
Sequence 1, Application US/08478689  
Patent No. 5614369  
GENERAL INFORMATION:  
APPLICANT: Wainwright, No. 5614369man R.  
APPLICANT: No. 5614369itsky, Thomas J.  
TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
TITLE OF INVENTION: Protein and Uses Thereof  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: United States of America  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,689  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/264,244  
FILING DATE: 22-JUNE-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1413.0010005  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
TELEX: 248636 SSK  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-478-689-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
Db 40 RRLKWKYKGF 50

RESULT 9  
US-08-478-689-3  
Sequence 3, Application US/08478689  
Patent No. 5614369  
GENERAL INFORMATION:  
APPLICANT: Wainwright, No. 5614369man R.  
APPLICANT: No. 5614369itsky, Thomas J.  
TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
TITLE OF INVENTION: Protein and Uses Thereof  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: United States of America  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,689  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/264,244  
FILING DATE: 22-JUNE-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1413.0010005  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
TELEX: 248636 SSK  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-478-689-3

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

US-08-456-112B-45

Query Match 100.0%; Score 64; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00017;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
|||||  
DB 1 KRLKWKYKGF 11

## RESULT 5

PCT-US96-10227-1  
; Sequence 1, Application PC/TUS9610227  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, JAMES A.  
; APPLICANT: SHEKHANI, MOHAMMED S.  
; APPLICANT: FIRCA, JOSEPH R.  
; APPLICANT: SCHATZ, ROBERT W.  
; APPLICANT: PUGH, CHARLES  
; APPLICANT: PANASIK JR., NICHOLAS  
; APPLICANT: STAFFORD, DOUGLAS C.  
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF SEPSIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10227  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/482,191  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/169,701  
; FILING DATE: 17-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/164,067  
; FILING DATE: 08-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/995,388  
; FILING DATE: 21-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: OPD-01280  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 1:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 36  
; OTHER INFORMATION: /note= "The amino acid at this  
; OTHER INFORMATION: location can be either Val or Ile."  
; FEATURE:  
; NAME/KEY: Peptide

; LOCATION: 102  
; OTHER INFORMATION: /note= "The amino acid at this  
; OTHER INFORMATION: location can be either Gln or Glu."  
PCT-US96-10227-1

Query Match 100.0%; Score 64; DB 3; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
|||||  
DB 41 KRLKWKYKGF 51

## RESULT 6

US-08-264-244-1  
; Sequence 1, Application US/08264244  
; Patent No. 5594113  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5594113man R.  
; APPLICANT: No. 5594113itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses Thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/264,244  
; FILING DATE: 22-JUNE-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1413.0010005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; TELEX: 248636 SSK  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-264-244-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
|||||  
DB 40 RRLKWKYKGF 50

## RESULT 7

US-08-264-244-3  
; Sequence 3, Application US/08264244  
; Patent No. 5594113  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5594113man R.  
; APPLICANT: No. 5594113itsky, Thomas J.



Db 41 KRLKWKYKGF 51  
 |||||

## RESULT 2

US-08-097-830E-28  
 ; Sequence 28, Application US/08097830E  
 ; Patent No. 5652211  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PORIO, Massimo  
 ; TITLE OF INVENTION: Peptides For Neutralizing The  
 ; TITLE OF INVENTION: Toxicity of Lipid A  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/097,830E  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-003  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 11 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-097-830E-28

Query Match 100.0%; Score 64; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00017;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 |||||

Db 1 KRLKWKYKGF 11

## RESULT 3

US-08-456-112B-28  
 ; Sequence 28, Application US/08456112B  
 ; Patent No. 5834430  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PORIO, Massimo  
 ; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: LEADING EDGE 486  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/456,112B  
 ; FILING DATE: May 31, 1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 11 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: circular  
 ; US-08-456-112B-28

Query Match 100.0%; Score 64; DB 2; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00017;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 |||||

Db 1 KRLKWKYKGF 11

## RESULT 4

US-08-456-112B-45  
 ; Sequence 45, Application US/08456112B  
 ; Patent No. 5834430  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PORIO, Massimo  
 ; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: LEADING EDGE 486  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/456,112B  
 ; FILING DATE: May 31, 1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 45:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 11 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: circular

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:45 ; Search time 82.37 Seconds  
(without alignments)  
1.318 Million cell updates/sec

Title: US-09-124-280A-28

Perfect score: 64

Sequence: 1 KRLWKYKGF 11

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	102	1	US-08-169-701-1
2	64	100.0	11	1	US-08-097-830E-28
3	64	100.0	11	2	US-08-456-112B-28
4	64	100.0	11	2	US-08-456-112B-45
5	64	100.0	102	3	PCT-US96-10227-1
6	61	95.3	101	1	US-08-264-244-1
7	61	95.3	105	1	US-08-264-244-3
8	61	95.3	101	1	US-08-478-689-1
9	61	95.3	105	1	US-08-478-689-3
10	61	95.3	101	1	US-08-476-940-1
11	61	95.3	105	1	US-08-476-940-3
12	61	95.3	101	1	US-08-704-872-1
13	61	95.3	105	1	US-08-704-872-3
14	61	95.3	101	2	US-08-577-464-2
15	59	92.2	11	2	US-08-366-953A-25
16	39	60.9	2004	1	US-08-375-709-15
17	39	60.9	2004	2	US-08-752-929-15
18	35	54.7	332	1	US-08-469-649-2
19	34	53.1	391	1	US-07-921-178A-2
20	34	53.1	25	1	US-07-921-178A-16
21	34	53.1	391	1	US-08-103-445-5
22	34	53.1	391	1	US-08-461-690B-5
23	34	53.1	548	1	US-08-468-700-37
24	34	53.1	548	2	US-08-645-971-5
25	34	53.1	548	2	US-08-468-220-35
26	34	53.1	548	2	US-08-468-698-35
27	34	53.1	399	2	US-08-501-003A-11
28	34	53.1	372	2	US-08-501-003A-12
29	34	53.1	391	2	US-08-501-003A-13
30	34	53.1	383	2	US-08-501-003A-14
31	34	53.1	398	2	US-08-501-003A-15
32	34	53.1	391	2	US-08-501-003A-16
33	34	53.1	548	3	PCT-US94-01553A-35
34	34	53.1	548	3	PCT-US95-10426-35
35	32	50.0	138	1	US-07-827-506B-4
36	32	50.0	510	1	US-08-097-829-2
37	32	50.0	426	1	US-08-336-583-2
38	32	50.0	510	2	US-08-577-403-2
39	32	50.0	376	2	US-08-709-923-4

40 32 50.0 426 3 PCT-US95-13795-2  
41 31 48.4 483 1 US-08-194-338-7  
42 31 48.4 997 1 US-08-232-540-1  
43 31 48.4 519 1 US-08-319-387-4  
44 31 48.4 1021 1 US-08-233-008A-2  
45 31 48.4 998 1 US-08-233-008A-6

#### ALIGNMENTS

RESULT 1  
US-08-169-701-1  
; Sequence 1, Application US/08169701  
; Patent No. 5545721  
; GENERAL INFORMATION:  
; APPLICANT: CARROLL, SEAN B.  
; APPLICANT: FIRCA, JOSEPH R.  
; APPLICANT: PUGH, CHARLES  
; APPLICANT: PADHYE, NISHA V.  
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF SEPSIS  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/169,701  
; FILING DATE: 17-DEC-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/995,388  
; FILING DATE: 21-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: OPHD-00748  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 36  
; OTHER INFORMATION: /note= "The amino acid at this  
; OTHER INFORMATION: location can be either Val or Ile."  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 102  
; OTHER INFORMATION: /note= "The amino acid at this  
; OTHER INFORMATION: location can be either Gln or Glu."  
; US-08-169-701-1

Query Match 100.0%; Score 64; DB 1; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLWKYKGF 11

CC binding molecule; and (c) removing the complex for the sample. The method  
 CC is useful removing endotoxins from contaminated blood. Endotoxins are  
 CC lipid, carbohydrate and protein complexes from bacteria, normally found  
 CC in their outer wall. They are shed from living gram negative bacteria and  
 CC released when they die leading to gram negative septic shock from  
 CC bacterial infection, especially after surgery. The protein used in the  
 CC method binds to endotoxins and neutralises them. The method removes  
 CC endotoxins from infected patients. Previous antibiotic treatment of  
 CC bacterial infection only kills the bacteria, still leaving the endotoxin  
 CC in the system. Antibodies with the same purpose have been developed, but  
 CC the binding protein used has a higher binding constant and as it has a  
 CC low molecular weight, it is less antigenic than other proteins used in  
 CC similar procedures. The method is better at removing endotoxin than  
 CC previous methods e.g. ultrafiltration, ion exchange, affinity  
 CC chromatography etc. which cannot distinguish between proteins and  
 CC endotoxins. 101 AA;

SQ Sequence 101 AA;

Query Match 95.3%; Score 61; DB 1; Length 101;  
 Best Local Similarity 90.9%; Pred. No. 0.0051;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 :|||||  
 Db 40 RRLKWKYKGF 50

# RESULT 14

W65112  
 ID W65112 standard; protein; 101 AA.

AC W65112;  
 DT 28-SEP-1998 (first entry)  
 DE Horseshoe crab ENP protein fragment.  
 KW Endotoxin neutralising protein; ENP; horseshoe crab; detection; biocide;  
 KW quantification; oil-in-water emulsion; coolant; metal cutting; aluminium;  
 KW steel rolling; metal forming; rust prevention; metal cleaner; antibiotic;  
 KW heat treating fluid; paint; adhesive; food production; waste water;  
 KW air quality; pharmaceutical; Gram positive bacteria; diagnostic.  
 OS Carcinocorpius rotundicauda.  
 PN W09821357-A1.

PD 22-MAY-1998.  
 PF 13-NOV-1997; U20556.  
 PR 10-NOV-1997; US-967038.  
 PR 13-NOV-1996; US-030594.  
 PA (ASCA-) ASSOC CAPE COD INC.  
 PI Novitsky TJ, Ridge RJ, Sloyer JL;  
 DR WPI; 98-297959/26.  
 PT Fluorescently labelled endotoxin neutralising proteins - used for  
 PT detection of endotoxins in e.g. metal working fluids, paints,  
 PT adhesives, foods or personal care, pharmaceutical or medical  
 PT products

PS Claim 12; Page 27; 42pp; English.  
 CC This sequence represents an endotoxin neutralising protein (ENP)  
 CC fragment from the horseshoe crab. This sequence is used in a method  
 CC which detects or quantifies endotoxins in samples such as oil-in-water  
 CC emulsions used as coolants for metal cutting, aluminium and steel  
 CC rolling, metal forming, rust preventatives, metal cleaners, metal heat  
 CC treating fluids, paints and adhesives, food products and food handling  
 CC surfaces, waste water and cooling tower water, personal care products,  
 CC indoor and outdoor air quality, pharmaceutical products, and medical  
 CC diagnostic products, e.g. blood, urine, CSF, or lavages. In addition the  
 CC method can be used to measure molecules specific for gram positive  
 CC bacteria and fungi, as well as biocides and antibiotics.  
 CC Sequence 101 AA;

Query Match 95.3%; Score 61; DB 1; Length 101;  
 Best Local Similarity 90.9%; Pred. No. 0.0051;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 :|||||

Db 40 RRLKWKYKGF 50

# RESULT 15

ID W65113 standard; Protein; 105 AA.

AC W65113;  
 DT 28-SEP-1998 (first entry)  
 DE Horseshoe crab recombinant ENP protein fragment.  
 KW Endotoxin neutralising protein; ENP; horseshoe crab; detection; biocide;  
 KW quantification; oil-in-water emulsion; coolant; metal cutting; aluminium;  
 KW steel rolling; metal forming; rust prevention; metal cleaner; antibiotic;  
 KW heat treating fluid; paint; adhesive; food production; waste water;  
 KW air quality; pharmaceutical; Gram positive bacteria; diagnostic.  
 OS Carcinocorpius rotundicauda.  
 PN W09821357-A1.

PD 22-MAY-1998.

PF 13-NOV-1997; U20556.

PR 10-NOV-1997; US-967038.

PR 13-NOV-1996; US-030594.

PA (ASCA-) ASSOC CAPE COD INC.

PI Novitsky TJ, Ridge RJ, Sloyer JL;

DR WPI; 98-297959/26.

DR N-PSDB; V35327.

PT Fluorescently labelled endotoxin neutralising proteins - used for  
 PT detection of endotoxins in e.g. metal working fluids, paints,  
 PT adhesives, foods or personal care, pharmaceutical or medical  
 PT products

PS Claim 13; Page 28; 42pp; English.

CC This sequence represents a recombinant endotoxin neutralising protein  
 CC (ENP) fragment from the horseshoe crab. This sequence is used in a method  
 CC which detects or quantifies endotoxins in samples such as oil-in-water  
 CC emulsions used as coolants for metal cutting, aluminium and steel  
 CC rolling, metal forming, rust preventatives, metal cleaners, metal heat  
 CC treating fluids, paints and adhesives, food products and food handling  
 CC surfaces, waste water and cooling tower water, personal care products,  
 CC indoor and outdoor air quality, pharmaceutical products, and medical  
 CC diagnostic products, e.g. blood, urine, CSF, or lavages. In addition the  
 CC method can be used to measure molecules specific for gram positive  
 CC bacteria and fungi, as well as biocides and antibiotics.  
 CC Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
 Best Local Similarity 90.9%; Pred. No. 0.0052;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 :|||||  
 Db 44 RRLKWKYKGF 54

Search completed: September 7, 1999, 22:49:59  
 Job time: 7880 sec

```

QY 1 KRLKWKYKGF 11
   :|||||
Db 44 RRLKWKYKGF 54

RESULT 11
W18220
ID W18220 standard; Protein; 101 AA.
AC W18220;
DT 31-JAN-1998 (first entry)
DE Limulus anti-lipopolysaccharide factor.
KW Anti-lipopolysaccharide factor; LPS; Limulus anti-LPS factor;
KW LALF; Gram-negative bacterium; infection; therapy; acne;
KW antimicrobial; preservative.
OS Limulus polyphemus.
QS Synthetic.
FH Key Location/Qualifiers
FT Domain 30..55
FT FT /note= "domain required for endotoxin binding and
FT neutralisation"
FT
PN W09723235-A1.
PD 03-JUL-1997.
PF 17-DEC-1996; U19882.
PR 22-DEC-1995; US-577464.
PA (MARI-) MARINE BIOLOGICAL LAB.
PI Wainwright NR.
DR WPI: 97-350781/32.
DR N-PSDB: 771720.
PT Anti-lipopolysaccharide factor proteins for use in cosmetic, skin or
PT hair preparations - useful to treat Gram-negative bacterial
PT infections, especially acne.
PS Example 1; Fig 1; 30pp; English.
CC This polypeptide comprises Limulus anti-lipopolysaccharide factor
CC (LALF) (W18220). It was expressed in Pichia host cells using a
CC synthetic gene (see T71720) that utilises Saccharomyces cerevisiae-
CC preferred codons. The yeast alpha-factor prepro leader sequence,
CC present on the expressed protein as an N-terminal fusion, targets
CC the protein for secretion into the medium and is cleaved off in the
CC process. A claimed pharmaceutical composition contains one or more
CC anti-lipopolysaccharide (anti-LPS) factor proteins and one or more
CC antibiotics. Also claimed is a preservative for use in cosmetic or
CC skin or hair preparations, comprising one or more anti-LPS factor
CC proteins. The composition is used to treat bacterial infections
CC (in conjunction with antimicrobials) and yeast infections. In
CC particular the composition is for treating skin acne, optionally in
CC conjunction with antibiotics.
SQ Sequence 101 AA;

Query Match 95.3%; Score 61; DB 1; Length 101;
Best Local Similarity 90.9%; Pred. No. 0.0051;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11
   :|||||
Db 40 RRLKWKYKGF 50

RESULT 12
W53876
ID W53876 standard; protein; 105 AA.
AC W53876;
DT 13-JUL-1998 (first entry)
DE L. polyphemus endotoxin.
KW Endotoxin contamination; horseshoe crab; gram negative septic shock;
KW bacterial infection.
OS Limulus polyphemus.
PN US5747455-A.
PD 05-MAY-1998.
PF 30-AUG-1996; 704872.
PR 15-MAY-1992; US-883457.
PR 23-JUN-1988; US-210575.
PR 16-FEB-1990; US-480957.
PR 16-MAY-1991; US-701501.
PR 22-JUN-1994; US-264244.
PR 30-AUG-1996; US-704872.
PA (ASCA-) ASSOC CAPE COD INC.
PI Novitsky TJ, Wainwright NR.
DR WPI: 98-285799/25.
PT Method for removing endotoxin contaminants - uses an endotoxin
PT binding protein from horseshoe crab
PS Disclosure; Fig 15; 29pp; English.
CC This sequence represents the Limulus polyphemus (horseshoe crab)
CC endotoxin. It can be used in the method of the invention for reducing
CC endotoxin contamination of a material containing it, comprising:
CC (a) contacting the material with the endotoxin binding molecule from
CC L. polyphemus; (b) forming a complex between the endotoxin and the
CC L. polyphemus; (b) forming a complex between the endotoxin and the

```

```

PR 16-FEB-1990; US-480957.
PR 16-MAY-1991; US-701501.
PR 22-JUN-1994; US-264244.
PR 30-AUG-1996; US-704872.
PA (ASCA-) ASSOC CAPE COD INC.
PI Novitsky TJ, Wainwright NR.
DR WPI: 98-285799/25.
DR N-PSDB: V23553.
PT Method for removing endotoxin contaminants - uses an endotoxin
PT binding protein from horseshoe crab
PS Claim 1; Fig 16; 29pp; English.
CC This sequence represents the Limulus polyphemus (horseshoe crab)
CC endotoxin. It can be used in the method of the invention for reducing
CC endotoxin contamination of a material containing it, comprising:
CC (a) contacting the material with the endotoxin binding molecule from
CC L. polyphemus; (b) forming a complex between the endotoxin and the
CC binding molecule; and (c) removing the complex for the sample. The method
CC is useful removing endotoxins from contaminated blood. Endotoxins are
CC lipid, carbohydrate and protein complexes from bacteria, normally found
CC in their outer wall. They are shed from living gram negative bacteria and
CC released when they die leading to gram negative septic shock from
CC bacterial infection, especially after surgery. The protein used in the
CC method binds to endotoxins and neutralises them. The method removes
CC endotoxins from infected patients. Previous antibiotic treatment of
CC bacterial infection only kills the bacteria, still leaving the endotoxin
CC in the system. Antibodies with the same purpose have been developed, but
CC the binding protein used has a higher binding constant and as it has a
CC low molecular weight, it is less antigenic than other proteins used in
CC similar procedures. The method is better at removing endotoxin than
CC previous methods e.g. ultrafiltration, ion exchange, affinity
CC chromatography etc. which cannot distinguish between proteins and
CC endotoxins.
SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;
Best Local Similarity 90.9%; Pred. No. 0.0052;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11
   :|||||
Db 44 RRLKWKYKGF 54

RESULT 13
W53877
ID W53877 standard; protein; 101 AA.
AC W53877;
DT 13-JUL-1998 (first entry)
DE L. polyphemus endotoxin.
KW Endotoxin contamination; horseshoe crab; gram negative septic shock;
KW bacterial infection.
OS Limulus polyphemus.
PN US5747455-A.
PD 05-MAY-1998.
PF 30-AUG-1996; 704872.
PR 15-MAY-1992; US-883457.
PR 23-JUN-1988; US-210575.
PR 16-FEB-1990; US-480957.
PR 16-MAY-1991; US-701501.
PR 22-JUN-1994; US-264244.
PR 30-AUG-1996; US-704872.
PA (ASCA-) ASSOC CAPE COD INC.
PI Novitsky TJ, Wainwright NR.
DR WPI: 98-285799/25.
PT Method for removing endotoxin contaminants - uses an endotoxin
PT binding protein from horseshoe crab
PS Disclosure; Fig 15; 29pp; English.
CC This sequence represents the Limulus polyphemus (horseshoe crab)
CC endotoxin. It can be used in the method of the invention for reducing
CC endotoxin contamination of a material containing it, comprising:
CC (a) contacting the material with the endotoxin binding molecule from
CC L. polyphemus; (b) forming a complex between the endotoxin and the

```

DR N-PSDB; T62063.  
PT New horse-shoe crab endotoxin-binding protein - useful for  
PT neutralising the effects of the endotoxin in an animal  
PS Claim 1; Columns 19-20; 29pp; English.  
CC The present sequence is a Limulus polyphemus (horseshoe crab)  
CC endotoxin binding protein modified for expression in a yeast host,  
CC which can be used to bind and neutralise endotoxin in vivo and in  
CC endotoxin assays. The protein can be used to treat septicemia,  
CC toxic shock, endotoxin related arthritis, gonorrhea, periodontal  
CC disease, spinal meningitis, amniotic fluid infection and gram  
CC negative bacterial infections in humans or other mammals. The  
CC protein has a very high binding constant, and its low molecular  
CC weight makes it likely to be less antigenic than anti-endotoxin  
CC antibodies.  
CC Fifty ng/kg of E. coli endotoxin was injected into 9 rabbits, and  
CC 15 minutes later 3 were injected with 5 microg of the protein, 3  
CC with 50 microg of the protein and 3 with PBS only. The control  
CC animals that received endotoxin and PBS only showed normal peak  
CC fever response 1 hour after toxin administration, while animals  
CC that received the protein showed a greatly reduced fever response  
CC proportional to the amount of protein administered.  
CC Sequence 105 AA;  
SQ

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0052;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
DB 44 RRLKWKYKGF 54

## RESULT 9

WI8326  
ID WI8326 standard; Protein; 105 AA.  
AC WI8326;  
DE 22-JUL-1997 (first entry)  
DE Endotoxin binding/neutralising protein.  
KW Horseshoe crab; L. polyphemus; endotoxin binding/neutralising protein;  
KW assay; water purity; cleanliness; pharmaceutical manufacture;  
KW kidney dialysis unit; Limulus amoebocyte lysate; LAL assay;  
KW septicemia; toxic shock; endotoxin related arthritis; gonorrhea;  
KW periodontal disease; spinal meningitis; amniotic fluid infection;  
KW septic shock; bacterial endotoxin.  
OS Limulus polyphemus.  
FH Key Location/Qualifiers  
FT peptide 1. .4  
FT protein /note= "N-terminal extension"  
FT protein 5. .105  
FT protein /note= "Endotoxin binding/neutralising protein"  
PN US5614369-A.  
PD 25-MAR-1997.  
PD 23-JUN-1988; 210575.  
PR 16-FEB-1990; US-210575.  
PR 16-FEB-1990; US-480957.  
PR 16-MAY-1991; US-701501.  
PR 15-MAY-1992; US-883457.  
PR 22-JUN-1994; US-264244.  
PR 07-JUN-1995; US-478689.  
PA (ASCA-) ASSOC CAPE COD INC.  
PI Novitsky TJ, Wainwright NR;  
DR WPI: 97-201472/18.  
DR N-PSDB; T68929.  
PT Assaying endotoxin using binding protein from horseshoe crab - and  
PT measuring quenching of tryptophan fluorescence caused by endotoxin  
PT binding  
PS Claim 1; Column 19-20; 29pp; English.  
CC This sequence represents the horseshoe crab (L. polyphemus) endotoxin  
CC binding/neutralising protein. This protein was used in the method of  
CC the invention for assaying endotoxin concentration. The method can  
CC be used to determine water purity and process cleanliness during  
CC pharmaceutical manufacture, to check kidney dialysis units and

CC generally wherever the Limulus amoebocyte lysate (LAL) assay is  
CC currently used. Also the protein, or a truncated version lacking the  
CC 1st 4 amino acids, can be used therapeutically to bind/neutralise  
CC endotoxin in vivo (e.g. in cases of septicemia, toxic shock, endotoxin  
CC related arthritis, gonorrhea, periodontal disease, spinal meningitis  
CC and amniotic fluid infection), and to remove endotoxin from solutions  
CC (e.g. extracorporeal treatment of septic shock). This protein has a  
CC very high binding constant for a wide range of bacterial endotoxins,  
CC and being of low molecular weight is likely to be less immunogenic  
CC than anti-endotoxin antibodies. The protein also avoids the multiple  
CC interferences associated with the conventional LAL assay.  
SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0052;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
DB 44 RRLKWKYKGF 54

## RESULT 10

WI9248  
ID WI9248 standard; Protein; 105 AA.  
AC WI9248;  
DT 18-AUG-1997 (first entry)  
DE Modified horseshoe crab endotoxin binding/neutralising protein.  
KW amino-terminal; tetrapeptide; yeast; restriction site; solid;  
KW resin; immobilisation; support; chromatographic; membrane;  
KW phase; latex microsphere; removal; solution; pharmaceutical;  
KW extracorporeal; treatment; septic shock; reagent; assay;  
KW fluorescence quenching; biosensor; determination;  
KW agglutination.  
OS Limulus polyphemus.  
FH Key Location/Qualifiers  
FT region 1. .4  
FT /note= "additional amino-terminal tetrapeptide"  
PN US5627266-A.  
PD 06-MAY-1997.  
PD 23-JUN-1988; 210575.  
PR 23-JUN-1988; US-210575.  
PR 16-FEB-1990; US-480957.  
PR 16-MAY-1991; US-701501.  
PR 15-MAY-1992; US-883457.  
PR 22-JUN-1994; US-264244.  
PR 07-JUN-1995; US-476940.  
PA (ASCA-) ASSOC CAPE COD INC.  
PI Novitsky TJ, Wainwright NR;  
DR WPI: 97-271365/24.  
DR N-PSDB; T69763.  
DT Immobilised endotoxin binding protein - for removal or determination  
PT of endotoxin  
PS Claim 1; Columns 19-20; 29pp; English.  
CC The present sequence is a modified Horseshoe crab endotoxin  
CC binding/neutralising protein having an additional amino-terminal  
CC tetrapeptide, the codons for which are yeast preferred codons  
CC possessing unique restriction sites. When immobilised on a solid  
CC support (preferably a chromatographic resin or membrane, or latex  
CC microspheres) the protein can be used to remove endotoxin from  
CC solutions, e.g. pharmaceutical solutions, for the extracorporeal  
CC treatment of septic shock, as a solid phase reagent in a fluorescence  
CC quenching assay for endotoxin, in biosensors for endotoxin  
CC determination or as an agglutination reagent for endotoxin  
CC determination.  
SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0052;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CC and has the additional coding sequence between bases 1-12 which encodes  
 CC the tetrapeptide "Glu-Ala-Glu-Ala" indicated in the features table.  
 CC This peptide is covalently linked to the amino terminal of the  
 CC endotoxin binding/neutralising protein.  
 SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
 Best Local Similarity 90.9%; Pred. No. 0.0052; 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 1; Mismatches 0;

QY 1 KRLKWKYKGRF 11  
 :|||||  
 Db 44 RRLKWKYKGRF 54

## RESULT 6

R29468 ID R29468 standard; Protein; 101 AA.  
 AC R29468; DT 15-APR-1993 (first entry)  
 DE Endotoxin binding/neutralising protein.  
 KW Endotoxin; binding; neutralising; protein; horseshoe; crab; lysate;  
 KW cellular debris; limulus; amebocyte; hypotonic shock; yeast; urea;  
 KW denaturant; guanidine hydrochloride; tetrapeptide.  
 OS Limulus polyphemus.  
 PS Claim 17; Fig 15; 67pp; English.  
 PD 26-NOV-1992; U03983.  
 PF 15-MAY-1992; US-701501.  
 PR 15-MAY-1991; US-701501.  
 PA (ASCA-) ASSOC CAPE COD INC.  
 PI Novitsky TJ, Wainwright NR;  
 DR WPI; 92-415712/50.  
 PT Use of endotoxin binding-neutralising protein isolated from  
 PT horseshoe crab - for reversing the effects of endotoxin in vivo,  
 PT and treating Gram negative bacterial infections, septicaemia,  
 PT toxic shock etc.  
 PS Claim 17; Fig 15; 67pp; English.  
 CC The sequence given is an endotoxin binding/neutralising protein  
 CC crab. This protein may be obtained from the cellular debris produced  
 CC during lysate production from Limulus amebocytes. The amebocytes are  
 CC lysed by hypotonic shock. The endotoxin binding/neutralising protein  
 CC was extracted from the cellular debris by using denaturants such as  
 CC urea and guanidine hydrochloride. The endotoxin binding/neutralisation  
 CC domain was found to occur between amino acids 30-55 of the sequence  
 CC given.  
 SQ Sequence 101 AA;

Query Match 95.3%; Score 61; DB 1; Length 101;  
 Best Local Similarity 90.9%; Pred. No. 0.0051; 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 1; Mismatches 0;

QY 1 KRLKWKYKGRF 11  
 :|||||  
 Db 40 RRLKWKYKGRF 50

## RESULT 7

R90148 ID R90148 standard; peptide; 22 AA.  
 AC R90148; DT 06-MAR-1996 (first entry)  
 DE LALF-(31-52) peptide.  
 KW lipopolysaccharide-binding peptide; LBP; LPS; antibacterial; sepsis;  
 KW limulus anti-LPS factor; LALF; endotoxin; amphipathic loop structure.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1..22  
 FT domain /note= "loop-stabilising bond"  
 FT 8..15  
 FT /note= "LPS-binding domain"

PN W09505393-A2.  
 PD 23-FEB-1995.  
 PF 18-AUG-1994; E02747.  
 PR 18-AUG-1993; US-108415.  
 PR 25-AUG-1993; US-111625.  
 PA (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.  
 PI Hoess A, Liddington RC;  
 DR WPI; 95-098719/13.  
 PT New lipopolysaccharide binding and neutralising peptide(s) - used  
 PT for treating or preventing LPS associated conditions or for  
 PT detecting or removing bacterial LPS.  
 PS Claim 11; Page 27; 46pp; English.  
 CC New LPS-binding peptides are claimed which have a domain of 8 amino  
 CC acids (AA) comprising (from N to C terminal): (1) a polar or positively  
 CC charged AA (pref. C,H,K,N,Q,R,S,T,W or Y); (2) a hydrophobic AA (pref.  
 CC A,F,H,I,L,M,V or W); (3) a basic AA (pref. H,K or R); (4) a hydrophobic  
 CC or positively charged AA (pref. A,F,H,I,K,L,M,R,V or W); (5) a  
 CC hydrophobic, polar or positively charged AA (pref. A,C,F,H,I,K,L,M,N,Q,  
 CC R,S,T,V,W or Y); (6) a positively charged AA (pref. K or R); (7) a  
 CC hydrophobic, polar or positively charged AA (pref. A,C,F,H,I,K,L,M,N,Q,  
 CC R,S,T,V,W or Y); and (8) a hydrophobic or positively charged AA (pref.  
 CC A,F,H,I,K,L,M,R,V or W). The peptide can alternatively be a  
 CC corresponding inverse sequence, or it can be an LPS-binding variation  
 CC of either sequence. The peptides have been designed by studying an  
 CC amphipathic loop structure found in the LPS-binding proteins limulus  
 CC anti-LPS factor (LALF), rabbit and human lipopolysaccharide-binding  
 CC protein (LBP) and human bactericidal/permeability-increasing protein  
 CC (BPI). The present sequence is a preferred example of the new peptides.  
 CC The peptides bind with high affinity to endotoxin and are useful for  
 CC prevention and treatment of Gram-negative and Gram-positive bacterial  
 CC sepsis and for the treatment of bacterial and fungal infections, as  
 CC well as for neutralising effects associated with heparin such as  
 CC anti-coagulation, angiogenesis, and growth factor-induced tumour and  
 CC endothelial cell proliferation.  
 SQ Sequence 22 AA;

Query Match 95.3%; Score 61; DB 1; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 0.0012; 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 1; Mismatches 0;

QY 1 KRLKWKYKGRF 11  
 :|||||  
 Db 10 RRLKWKYKGRF 20

## RESULT 8

W13262 ID W13262 standard; Protein; 105 AA.  
 AC W13262; DT 29-MAY-1997 (first entry)  
 DE Modified Limulus polyphemus endotoxin binding protein.  
 KW Neutralisation; In vivo; assay; horseshoe crab; endotoxin; binding protein;  
 KW neutralisation; In vivo; assay; treatment; septicaemia; yeast;  
 KW toxic shock; endotoxin related arthritis; gonorrhea; expression;  
 KW periodontal disease; spinal meningitis; amniotic fluid infection;  
 KW gram negative; bacterium; bacteria; infection; host.  
 OS Limulus polyphemus.  
 FH Key Location/Qualifiers  
 FT peptide 1..4  
 FT /note= "tetrapeptide to enable expression as  
 FT glycoprotein in yeast host"

US5594113-A.  
 PD 14-JAN-1997.  
 PF 23-JUN-1988; 210575.  
 PR 23-JUN-1988; US-210575.  
 PR 16-FEB-1990; US-480957.  
 PR 16-MAY-1991; US-701501.  
 PR 15-MAY-1992; US-883457.  
 PR 22-JUN-1994; US-264244.  
 PA (ASCA-) ASSOC CAPE COD INC.  
 PI Novitsky TJ, Wainwright NR;  
 DR WPI; 97-107149/10.

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 64; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00021;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGRF 11  
 |||||  
 Db 1 KRLKWKYKGRF 11

## RESULT 3

ID W21616 standard; peptide; 11 AA.  
 AC W21616;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #28.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 23-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence, reduces dose of antibiotic required  
 CS Claim 32; Page 27; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when these are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 64; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00021;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGRF 11  
 |||||  
 Db 1 KRLKWKYKGRF 11

## RESULT 4

ID W10350 standard; peptide; 102 AA.  
 AC W10350;  
 DT 22-SEP-1997 (first entry)  
 DE Limulus antilipopolysaccharide factor.  
 DE LALF; antimicrobial; immunoglobulin; surface binding ligand;  
 KW haptens; sepsis; influenza; viremia; fungemia; neurology; cancer;  
 KW endocrinology; antibiotic; antibody; antigenic.  
 OS Limulus polyphemus.  
 FH Key Location/Qualifiers  
 FT misc\_difference 36  
 FT /label= Val, Ile  
 PN W09640251-A1.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; U10227.

PR 07-JUN-1995; US-482191.  
 PA (OPHI-) OPHIDIAN PHARM INC.  
 PI Firca JR, Panasiak N, Pugh C, Schatz RW, Shekhani MS;  
 PI Stafford D, Williams JA;  
 DR WPI: 97-077224/07.  
 PT Antimicrobial compns. for diagnosis and therapy - comprising  
 PT microbial surface binding ligand-spacer-hapten, non-covalently  
 PT associated with immunoglobulin  
 PS Example 38; Fig 13; 227pp; English.  
 CC The present sequence is Limulus antilipopolysaccharide factor  
 CC (LALF), which is a single chain peptide known to bind and neutralise  
 CC endotoxin. This is conjugated to human immunoglobulin IgG to produce  
 CC an antibody, which is a conjugated antibody and antibiotic. Making  
 CC a conjugate, comprises covalently linking a surface-binding ligand  
 CC (SBL) to a hapten via a spacer to form a ligand-spacer-hapten (LSH)  
 CC compound, and reacting the compound with immunoglobulin (Ig) under  
 CC conditions that allow for non-covalent binding of the Ig to the  
 CC hapten of the compound. The conjugate can be used for the diagnosis,  
 CC prevention and treatment of microbial infections, e.g. sepsis,  
 CC influenza, viremia or fungemia. They can also be used in neurology,  
 CC cancer and endocrinology, where Ig targeting can provide desired  
 CC therapeutic effects. The methods for producing a conjugate can use  
 CC small molecule targeting ligands to efficiently direct otherwise  
 CC unreactive Ig to microbial targets. This targeting allows Ig to react  
 CC with microbial structures that may not normally be accessible to Ig,  
 CC or are incapable of stimulating antibody production.  
 SQ Sequence 102 AA;

Query Match 100.0%; Score 64; DB 1; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 0.0017;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGRF 11  
 |||||  
 Db 41 KRLKWKYKGRF 51

## RESULT 5

ID R29467 standard; Protein; 105 AA.  
 AC R29467;  
 DT 15-APR-1993 (first entry)  
 DE Modified endotoxin binding/neutralising protein.  
 KW Endotoxin; binding; neutralising; protein; horseshoe; crab; lysate;  
 KW cellular debris; Limulus; amoebocyte; hypotonic shock; yeast; urea;  
 KW denaturant; guanidine hydrochloride; tetrapeptide.  
 OS Limulus polyphemus.  
 FH Key Location/Qualifiers  
 FT peptide 1..4  
 FT /note= "Amino terminal modifying peptide"

WO9220715-A.  
 DT 26-NOV-1992.  
 PF 15-MAY-1992; U03983.  
 PR 16-MAY-1991; US-701501.  
 PA (ASCA-) ASSOC CAPE COD INC.  
 PI Novitsky TJ, Wainwright NR;  
 DR WPI: 92-415712/50.  
 DR N-PSDB; Q31793.

PT Use of endotoxin binding-neutralising protein isolated from  
 PT horseshoe crab - for reversing the effects of endotoxin in vivo,  
 PT and treating Gram negative bacterial infections, septicaemia,  
 PT toxic shock etc.  
 PS Claim 18; Fig 16; 67pp; English.  
 CC The sequence given is an endotoxin binding/neutralising protein  
 CC which may be isolated from any of the four known species of horseshoe  
 CC crab. This protein may be obtained from the cellular debris produced  
 CC during lysate production from Limulus amoebocytes. The amoebocytes are  
 CC lysed by hypotonic shock. The endotoxin binding/neutralising protein  
 CC was extracted from the cellular debris by using denaturants such as  
 CC urea and guanidine hydrochloride. The endotoxin binding/neutralisation  
 CC domain was found to occur between amino acids 34-59 of the sequence  
 CC given. The gene encoding this protein was isolated from a yeast host

GenCore version 4.5  
Copyright (C) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:49:57 ; Search time 135.78 Seconds  
(without alignments)  
1.919 Million cell updates/sec

Title: US-09-124-280A-28  
Perfect score: 64  
Sequence: 1 KRLKWKYKGF 11

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	64	100.0	11	1 R71799	Peptide neutralisi
2	64	100.0	11	1 W21633	Antibiotic potenti
3	64	100.0	11	1 W21616	Antibiotic potenti
4	64	100.0	102	1 W10350	Limulus antilipopo
5	61	95.3	105	1 R29467	Modified endotoxin
6	61	95.3	101	1 R29468	Endotoxin binding/
7	61	95.3	22	1 R90148	LALF-(31-52) pepti
8	61	95.3	105	1 W13262	Modified limulus p
9	61	95.3	105	1 W18326	Endotoxin binding/
10	61	95.3	105	1 W19248	Modified horseshoe
11	61	95.3	101	1 W18220	Limulus anti-lipop
12	61	95.3	105	1 W53876	L. polyphemus endo
13	61	95.3	101	1 W53877	L. polyphemus endo
14	61	95.3	101	1 W65112	Horseshoe crab ENP
15	61	95.3	105	1 W65113	Horseshoe crab rec
16	61	95.3	22	1 W99736	Limulus anti-lipop
17	43	67.2	237	1 W74827	Human secreted pro
18	42	65.6	660	1 P82921	B subunit of human
19	39	60.9	2004	1 R99464	Biosynthetic enzym
20	39	60.9	2004	1 W37052	S. putrefaciens EP
21	38	59.4	142	1 W36904	CD47 extracellular
22	37	57.8	1784	1 R05898	Gene product of fl
23	37	57.8	1784	1 R94427	FelV F6A provirus
24	37	57.8	243	1 W62715	Streptococcus pneu
25	35	54.7	332	1 R03167	Amino acid sequenc
26	35	54.7	954	1 W19752	Yeast inhibitor of
27	35	54.7	472	1 W86327	Kidney injury asso
28	34	53.9	813	1 W84307	A human CDC28-#46
29	34	53.1	372	1 R61334	Human potassium ch
30	34	53.1	391	1 R61335	Human potassium ch
31	34	53.1	266	1 R61333	Human potassium ch
32	34	53.1	372	1 R61332	Human potassium ch
33	34	53.1	389	1 R61331	Human potassium ch
34	34	53.1	391	1 R63233	ATP-sensitive K ch
35	34	53.1	372	1 R89175	Renal ATP-dependen
36	34	53.1	372	1 R89176	Renal ATP-dependen
37	34	53.1	778	1 R89174	Renal ATP-dependen
38	34	53.1	372	1 R89183	Renal ATP-dependen
39	34	53.1	391	1 R89182	Renal ATP-dependen
40	34	53.1	450	1 R99426	Murine lymphocyte
41	34	53.1	450	1 R99427	Human lymphocyte s
42	34	53.1	568	1 W19395	Human calpastatin
43	34	53.1	548	1 W19744	B. stearotheomphi

44 34 53.1 451 1 W38426 Human multiple mye  
45 34 53.1 873 1 W49031 Human PrtII-like su

ALIGNMENTS

RESULT 1

R71799  
ID R71799 standard; peptide; 11 AA.  
AC R71799;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN WO9503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids  
PT amino acids  
PS Claim 30; Page 22; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which include units of formula: (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new peptides.  
SQ Sequence 11 AA;

Query Match 100.0%; Score 64; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. NO. 0.00021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
|||||||  
Db 1 KRLKWKYKGF 11

RESULT 2

W21633  
ID W21633 standard; peptide; 11 AA.  
AC W21633;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #45.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide; permeability; outer bacterial membrane.  
OS Synthetic.  
PN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic required  
PS Claim 49; Page 29; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to potentiate the activity of an antibiotic when they are co-administered with the antibiotic. Compositions containing these peptides are used to treat or prevent microbial infections. These peptides bind to





RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA GEISEL C.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; U39472; AAA80134.1; -.  
 DR PFAM; PF00067; p450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.  
 FT BINDING 471 471 HEME (BY SIMILARITY).  
 SQ SEQUENCE 548 AA; 63665 MW; C4F8339E CRC32;

Query Match 63.6%; Score 35; DB 5; Length 548;  
 Best Local Similarity 62.5%; Pred. No. 92;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRF 8  
 ||:|:|  
 Db 146 KWREQRF 153

RESULT 14  
 O97983 PRELIMINARY; PRT; 74 AA.  
 ID O97983;  
 AC O97983;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE MHC CLASS I ALPHA 2 CHAIN (FRAGMENT).  
 OS Trionyx sinensis (Chinese softshell turtle) (Pelodiscus sinensis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Testudines;  
 OC Cryptodira; Trionychoidae; Trionychidae; Trionyx.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA XIA C.;  
 RT "MHC Class I Alpha 2 Domain of Soft-shelled turtle (Trionyx  
 RT sinensis).";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB022885; CAB19842.1; -.  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 74 74  
 SQ SEQUENCE 74 AA; 8670 MW; FCB19E68 CRC32;

Query Match 63.6%; Score 35; DB 7; Length 74;  
 Best Local Similarity 50.0%; Pred. No. 14;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRF 10  
 ||:|:|  
 Db 50 KWAEAF 59

RESULT 15  
 O92RT1 PRELIMINARY; PRT; 589 AA.  
 ID O92RT1;  
 AC O92RT1;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE GAMMA RESPONSE I PROTEIN.  
 DN GR I.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eudicotyledons; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA DEVEAUX Y., KAZMAIER M.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ131708; CAA10484.1; -.  
 SQ SEQUENCE 589 AA; 67472 MW; 798F6BB8 CRC32;

Query Match 63.6%; Score 35; DB 10; Length 589;  
 Best Local Similarity 66.7%; Pred. No. 98;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAQKRF 9  
 ||:|:|  
 Db 331 KWKSQHTFL 339

Search completed: September 7, 1999, 22:47:37  
 Job time: 7957 sec

RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA FUTON B., WOHLDMANN P.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF016446; AAC24164.1; -.  
 SQ SEQUENCE 321 AA; 37579 MW; BB77BA41 CRC32;

Query Match 65.6%; Score 36; DB 5; Length 321;  
 Best Local Similarity 66.7%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WKAKRFLK 10  
 |||||  
 Db 232 WKVKQLLK 240

## RESULT 10

ID O67695 PRELIMINARY; PRT; 373 AA.  
 AC O67695;  
 DT 01-AUG-1998 (TRENBLrel. 07, Created)  
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE PEPTIDE CHAIN RELEASE FACTOR RF-2.  
 GN PRFB.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE: 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus";  
 RL Nature 392:353-358(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE000758; AAC07656.1; -.  
 DR PFAM; PF00472; RF-1; 1.  
 SQ SEQUENCE 373 AA; 43309 MW; 921E417E CRC32;

Query Match 63.6%; Score 35; DB 2; Length 373;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQK 6  
 |||||  
 Db 361 KWKAQK 366

## RESULT 11

ID Q13647 PRELIMINARY; PRT; 568 AA.  
 AC Q13647;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last annotation update)  
 DE TESTIS SPECIFIC BASIC PROTEIN.

GN TSBP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA O'HERN P.A., YAVETZ H., MOY T., YAVETZ B., LIANG Z.G., WANG G.Y.,  
 RA GOLDBERG E.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U60665; AAB02975.1; -.  
 SQ SEQUENCE 568 AA; 63319 MW; 2A2CBEE0 CRC32;

Query Match 63.6%; Score 35; DB 4; Length 568;  
 Best Local Similarity 50.0%; Pred. No. 95;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10  
 |||||  
 Db 557 KWKSDFK 566

## RESULT 12

ID Q92538 PRELIMINARY; PRT; 1691 AA.  
 AC Q92538;  
 DT 01-FEB-1997 (TRENBLrel. 02, Created)  
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE MYELOBLAST KIAA0248 (FRAGMENT).  
 GN KIAA0248.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BONE MARROW;  
 RX MEDLINE: 97191544.  
 RA NAGASE T., SEKI N., ISHIKAWA K., OHIRA M., KAWARABAYASI Y., OHARA O.,  
 RA TANAKA A., KOTANI H., MIYAJIMA N., NOMURA N.;  
 RT "Prediction of the coding sequences of unidentified human genes. VI.  
 The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by  
 RT analysis of cDNA clones from cell line KG-1 and brain.";  
 RL DNA Res. 3:321-329(1996).  
 DR EMBL; D87435; BAA13379.1; -.  
 DR PFAM; PF01369; Sec7; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 1691 AA; 187806 MW; 8D8EC432 CRC32;

Query Match 63.6%; Score 35; DB 4; Length 1691;  
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAQK 7  
 |||||  
 Db 58 KWKAQK 64

## RESULT 13

ID Q27465 PRELIMINARY; PRT; 548 AA.  
 AC Q27465;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE SIMILAR TO CYTOCHROME P450 (EC 1.14.14.1).  
 GN B0304.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.

OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL: Z99710; CAB6866.1; -;  
 SQ SEQUENCE 662 AA; 74771 MW; DD0EADA7 CRC32;

Query Match 67.3%; Score 37; DB 5; Length 662;  
 Best Local Similarity 66.7%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAKQKFLK 10  
 ||| :|||  
 DB 497 WKATEKFLK 505

RESULT 7  
 O29641 ID O29641 PRELIMINARY; PRT; 93 AA.  
 AC O29641;  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TREMblrel. 07, Last annotation update)  
 DE HYPOTHETICAL 11.1 KD PROTEIN.  
 GN AF0614.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 OC Archaeoglobus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE: 98049343.  
 RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
 RA KETCHUM K.A., DODSON R.J., WINN M., HICKEY E.K., PETERSON J.D.,  
 RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDIS N.C.,  
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
 RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLOBEK A., ZHOU L.,  
 RA OVERBEER K., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTERBACK T.,  
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,  
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESSE C.R.,  
 RA VENTER J.C.;  
 RT "The complete genome sequence of the hyperthermophilic,  
 RT sulphate-reducing archaeon Archaeoglobus fulgidus."  
 RL Nature 390:364-370(1997).  
 DR EMBL: AE001062; AAB90629.1; -;  
 DR TIGR: AF0614; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 93 AA; 11148 MW; 0961504C CRC32;

Query Match 65.5%; Score 36; DB 1; Length 93;

Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KWKAQKRFK 10  
 :|| |||::  
 DB 75 EWKMSKRFIR 84  
 RESULT 8  
 O12154 ID O12154 PRELIMINARY; PRT; 354 AA.  
 AC O12154;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMblrel. 01, Last annotation update)  
 DE ORF YD1100C FROM LEFT ARM OF CHROMOSOME IV.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FY1679;  
 RA BOSKOVIC J., SAIZ J.E., SOLER-MIRA A., GARCIA-CANTALEJO J.,  
 RA REVUELTA J.L., JIMINEZ A., BALLESTA J.P.G., DEL REY F., REMACHA M.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA BALLESTA J.P.G., REMACHA M., SOLER-MIRA A., JIMENEZ A.,  
 RA GARCIA-CANTALEJO J.M., BOSKOVIC J., DEL REY F., REVUELTA J.L.,  
 RA BUITRAGO M.J., SANZ J.E.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: X95644; CAA64913.1; -;  
 DR EMBL: Z74148; CAA98667.1; -;  
 SQ SEQUENCE 354 AA; 39353 MW; 63AB1475 CRC32;

Query Match 65.5%; Score 36; DB 3; Length 354;  
 Best Local Similarity 55.6%; Pred. No. 40;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 9  
 :|| |||::  
 DB 291 RWMQKRYL 299

RESULT 9  
 O16430 ID O16430 PRELIMINARY; PRT; 321 AA.  
 AC O16430;  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)  
 DE C02E7.5 PROTEIN.  
 GN C02E7.5  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,

Query Match 67.3% Score 37; DB 3; Length 1955;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 WKAQKRFK 10  
 I : I I I I I  
 Db 1446 WRATKRFK 1454

RESULT 3  
 Q15312 PRELIMINARY; PRT; 1324 AA.  
 ID Q15312  
 AC Q15312;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE R KAPPA B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA NIETERS A., BOURMESTER T., SCHEIDEREIT C.;  
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92135142.  
 RA ADAMS B.S., LEUNG K.Y., HANLEY E.W., NABEL G.J.;  
 RT "Cloning of R kappa B, a novel DNA-binding protein that recognizes  
 the interleukin-2 receptor alpha chain kappa B site."  
 RL New Biol. 3:1063-1073(1991).  
 DR EMBL: X80878; CAA56846.1; -. Q -> G (IN REF. 2).  
 FT CONFLICT 97 97  
 SQ SEQUENCE 1324 AA; 141859 MW; E64355AD CRC32;

Query Match 67.3% Score 37; DB 4; Length 1324;  
 Best Local Similarity 50.08; Pred. No. 91;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10  
 I : I I I I I  
 Db 198 EWRTQRYLK 207

RESULT 4  
 O62137 PRELIMINARY; PRT; 679 AA.  
 ID O62137  
 AC O62137;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
 DE F08A8.2 PROTEIN.  
 GN F08A8.2  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 CC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z99710; CAB16865.1; -.  
 SQ SEQUENCE 679 AA; 77310 MW; E86F3685 CRC32;

RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z99710; CAB16864.1; -.  
 SQ SEQUENCE 679 AA; 77310 MW; E86F3685 CRC32;

Query Match 67.3% Score 37; DB 5; Length 679;  
 Best Local Similarity 66.7%; Pred. No. 49;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAQKRFK 10  
 I : I I I I I  
 Db 514 WKATEKFLK 522

RESULT 5  
 O62138 PRELIMINARY; PRT; 660 AA.  
 ID O62138  
 AC O62138;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
 DE F08A8.3 PROTEIN.  
 GN F08A8.3  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 CC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z99710; CAB16865.1; -.  
 SQ SEQUENCE 660 AA; 74781 MW; C24B6FFE CRC32;

Query Match 67.3% Score 37; DB 5; Length 660;  
 Best Local Similarity 66.7%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAQKRFK 10  
 I : I I I I I  
 Db 495 WKATEKFLK 503

RESULT 6  
 O62139 PRELIMINARY; PRT; 662 AA.  
 ID O62139  
 AC O62139;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
 DE F08A8.4 PROTEIN.  
 GN F08A8.4

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:36 ; Search time 160.2 Seconds  
(without alignments)  
3.842 Million cell updates/sec

Title: US-09-124-280A-27

Perfect score: 55

Sequence: 1 WKAKRFLK 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10:\*\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	69.1	469	3	059710 schizosacch
2	37	67.3	1955	3	074475 schizosacch
3	37	67.3	1324	4	015312 homo sapien
4	37	67.3	679	5	062137 caenorhabdi
5	37	67.3	660	5	062138 caenorhabdi
6	37	67.3	682	5	062139 caenorhabdi
7	36	65.5	93	1	029641 archaeoglob
8	36	65.5	354	3	012154 saccharomyc
9	36	65.5	321	5	016430 caenorhabdi
10	35	63.6	373	2	067695 aquifex aeo
11	35	63.6	568	4	013647 homo sapien
12	35	63.6	1691	4	092538 homo sapien
13	35	63.6	548	5	027465 caenorhabdi
14	35	63.6	74	7	097983 trionyx sin
15	35	63.6	589	10	092RT1 arabidopsis
16	35	63.6	2344	12	086114 rabbit hemo
17	35	63.6	2344	12	089273 rabbit hemo
18	35	63.6	2344	12	086117 rabbit hemo
19	35	63.6	1080	12	083895 ovine adeno
20	35	63.6	2344	12	086119 rabbit hemo
21	34	61.8	141	1	050529 methanobact
22	34	61.8	1732	2	051817 porphyromon
23	34	61.8	1732	2	052050 porphyromon
24	34	61.8	384	2	052050 porphyromon
25	34	61.8	1732	2	072197 porphyromon
26	34	61.8	1732	2	007442 porphyromon
27	34	61.8	292	2	025504 helicobacte
28	34	61.8	369	2	051304 borrelia bu
29	34	61.8	131	2	067614 aquifex aeo

30 34 61.8 456 2 082855 vibrio para  
31 34 61.8 1723 2 072194 porphyromon  
32 34 61.8 156 2 074448 synchocyst  
33 34 61.8 329 5 016431 caenorhabdi  
34 34 61.8 646 5 045076 caenorhabdi  
35 34 61.8 1038 5 061199 caenorhabdi  
36 34 61.8 660 5 017248 boophilus m  
37 34 61.8 680 5 094196 caenorhabdi  
38 34 61.8 364 5 002303 caenorhabdi  
39 34 61.8 462 10 004622 arabidopsis  
40 34 61.8 467 10 023384 arabidopsis  
41 33 60.0 337 2 067278 aquifex aeo  
42 33 60.0 4450 2 044928 bacillus br  
43 33 60.0 2273 3 042823 saccharomyc  
44 33 60.0 490 3 060016 schizosacch  
45 33 60.0 467 11 097431 mus musculu

## ALIGNMENTS

RESULT 1  
O59710 PRELIMINARY: PRT: 469 AA.  
ID O59710  
AC O59710;  
DT 01-AUG-1998 (T-EMBLrel. 07, Created)  
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)  
DT 01-AUG-1998 (T-EMBLrel. 07, Last annotation update)  
DE FERREDOXIN-NADP+ REDUCTASE.  
GN SPBC398.01C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H-;  
RA LYNE M., RAJANDREAM M.A., BARRELL B.G., BECK A., REINHARDT R.,  
RA POHL T;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL022244; CAA18290.1; -  
SQ SEQUENCE 469 AA; 52715 MW; 54915B59 CRC32;

Query Match 69.1%; Score 38; DB 3; Length 469;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 WKAKRFLK 10  
Db 402 WSKKFLK 410  
||::|||  
||::|||

RESULT 2  
O74475 PRELIMINARY: PRT: 1955 AA.  
ID O74475  
AC O74475;  
DT 01-NOV-1998 (T-EMBLrel. 08, Created)  
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (T-EMBLrel. 08, Last annotation update)  
DE PUTATIVE 1.3-BETA-GLUCAN SYNTHASE COMPONENT.  
GN SPCG1840.02C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H-;  
RA RIEGER M., WOOD V., RAJANDREAM M.A., BARRELL B.G.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL031179; CAA20125.1; -  
SQ SEQUENCE 1955 AA; 225185 MW; C592A33D CRC32;

```

CC -!- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL
CC LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER
CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS
CC TO INTERACT WITH THE CD14 RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE BPI/CE1P/LBP/PLTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L32132; G510715; -
CC PROSITE: PS00400; LBP_BPI_CETP; 1.
CC PFAM: PF01273; LBP_BPI_CETP; 1.
CC HSP: P17213; LBPI.
CC KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
CC FT SIGNAL 1 25 POTENTIAL.
CC FT CHAIN 26 481 LIPOPOLYSACCHARIDE-BINDING PROTEIN.
CC FT CARBOHYD 300 300 POTENTIAL.
CC FT CARBOHYD 355 355 POTENTIAL.
CC SQ SEQUENCE 481 AA; 53600 MW; 0BF55E7D CRC32;

```

```

Query Match 63.6%; Score 35; DB 1; Length 481;
Best Local Similarity 50.0%; Pred.No. 25;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 KWKAKQKFLK 10
   ||| :: |||
Db 115 KWKVRRSFVK 124

```

```

Search completed: September 7, 1999, 23:59:08
Job time: 520 sec

```

CC -!- SIMILARITY: CONTAINS 1 CBS DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: M88344; G206597; -;  
 CC EMBL: M88346; G206600; -;  
 CC EMBL: D01098; D1001347; -;  
 CC PIR: JX0145; JX0145;  
 CC DR PIR: A42790; A42790.  
 CC DR PROSITE: PS00901; CYS-SYNTHASE; 1.  
 CC DR PFAM: PF00291; S-T.dehydratase; 1.  
 CC DR PFAM: PF00571; CBS; 1.  
 CC KW LYASE: CYSTEINE BIOSYNTHESIS; PYRIDOXAL PHOSPHATE; HEME; CBS DOMAIN;  
 KW ALTERNATIVE SPLICING.  
 FT INIT\_MET 0 0  
 FT BINDING 115 115 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT DOMAIN 411 464 CBS.  
 FT VARSPLIC 513 513 S->Y (IN TYPE III).  
 FT VARSPLIC 514 527 MISSING (IN TYPE III).  
 FT VARSPLIC 449 477 AILGMVTLGNMLSLLAGKVRPSDEVCKV ->  
 FT LKQSDICHTPKRHLIQAHGLRKVPDTEA (IN TYPE  
 FT IV).  
 FT VARSPLIC 478 560 MISSING (IN TYPE III).  
 FT CONFLICT 414 414 L -> P (IN REF. 2).  
 FT SEQUENCE 560 AA; 61323 MW; 6D972105 CRC32;

Query Match 63.6%; Score 35; DB 1; Length 560;  
 Best Local Similarity 60.0%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KWKAQRFLK 10  
 II II I I I  
 Db 385 KMWLQGFEMK 394

RESULT 14  
 ID CCAL\_DROME STANDARD; PRT; 1851 AA.  
 AC P91645;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DE CALCIUM CHANNEL ALPHA-1 SUBUNIT (CACOPHONY PROTEIN) (NIGHTBLIND A  
 DE PROTEIN) (NO-ON-TRANSIENT B PROTEIN).  
 GN NBA OR DMCA1A OR CAC OR NONB.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97141514.  
 RA SMITH L.A., WANG X.J., PEIXOTO A.A., NEUMANN E.K., HALL L.M.,  
 RA HALL J.C.;  
 RT "A Drosophila calcium channel alpha subunit gene maps to a genetic  
 RT locus associated with behavioral and visual defects.";  
 RL J. NEUROSCI. 16:7868-7879(1996).  
 CC -!- FUNCTION: CALCIUM CHANNEL. WHEN MUTATED IT IS INVOLVED IN THE  
 CC GENERATION OF BEHAVIORAL, PHYSIOLOGICAL, AND LETHAL PHENOTYPES.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- DEVELOPMENTAL STAGE: ITS EXPRESSION PEAKS IN THE FIRST LARVAL  
 CC INSTAR, MIDPUPAL, AND LATE PUPAL STAGES. IN LATE-STAGE EMBRYOS, IT  
 CC IS EXPRESSED PREFERENTIALLY IN THE NERVOUS SYSTEM.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: U55776; G1737064; -;  
 CC FLYBASE; FBgn0005563; nba.  
 CC DR PFAM: PF00520; ION\_TRANS; 4.  
 CC KW IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;  
 KW CALCIUM CHANNEL; GLYCOPROTEIN.  
 FT TRANSMEM 37 57 POTENTIAL.  
 FT TRANSMEM 79 99 POTENTIAL.  
 FT TRANSMEM 108 128 POTENTIAL.  
 FT TRANSMEM 168 188 POTENTIAL.  
 FT TRANSMEM 257 277 POTENTIAL.  
 FT TRANSMEM 291 311 POTENTIAL.  
 FT TRANSMEM 440 460 POTENTIAL.  
 FT TRANSMEM 470 490 POTENTIAL.  
 FT TRANSMEM 511 531 POTENTIAL.  
 FT TRANSMEM 566 586 POTENTIAL.  
 FT TRANSMEM 607 627 POTENTIAL.  
 FT TRANSMEM 647 667 POTENTIAL.  
 FT TRANSMEM 768 788 POTENTIAL.  
 FT TRANSMEM 808 828 POTENTIAL.  
 FT TRANSMEM 837 857 POTENTIAL.  
 FT TRANSMEM 906 926 POTENTIAL.  
 FT TRANSMEM 975 995 POTENTIAL.  
 FT TRANSMEM 1021 1041 POTENTIAL.  
 FT TRANSMEM 1091 1111 POTENTIAL.  
 FT TRANSMEM 1130 1150 POTENTIAL.  
 FT TRANSMEM 1161 1181 POTENTIAL.  
 FT TRANSMEM 1223 1243 POTENTIAL.  
 FT TRANSMEM 1271 1291 POTENTIAL.  
 FT TRANSMEM 1317 1337 POTENTIAL.  
 FT CARBOHYD 234 234 POTENTIAL.  
 FT CARBOHYD 235 235 POTENTIAL.  
 FT CARBOHYD 865 865 POTENTIAL.  
 FT CARBOHYD 1430 1430 POTENTIAL.  
 FT SEQUENCE 1851 AA; 212168 MW; 601976A1 CRC32;

Query Match 63.6%; Score 35; DB 1; Length 1851;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAQKRF 8  
 I I I I I I I  
 Db 424 WRAEKRF 430

RESULT 15

ID LBP\_RAT STANDARD; PRT; 481 AA.  
 AC Q63313;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).  
 GN LBP.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=LIVER;  
 RX MEDLINE; 94292804.  
 RA SU G.L., FREESWICK P.D., GELLER D.A., WANG Q., SHAPIRO R.A., WAN Y.H.,  
 RA BILLAR T.R., TWEARDY D.J., SIMMONS R.L., WANG S.C.;  
 RT "Molecular cloning, characterization, and tissue distribution of rat  
 RT lipopolysaccharide binding protein. Evidence for extrahepatic  
 RT expression.";  
 RL J. IMMUNOL. 153:743-752(1994).



```

Db 142 RWEAKPFL 150
RESULT 11
YHY4_YEAST
ID YHY4_YEAST STANDARD; PRT; 571 AA.
AC P38871;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 65.8 KD PROTEIN IN GNDI-1K1 INTERGENIC REGION.
GN YHR184W.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RX MEDLINE; 94378003.
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
RA DU Z., FAVELLO A., FULTON L., GATTUNG S., GEISEL C., KIRSTEN J.,
RA KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
RA LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
RA NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASKIS E., VAUGHAN K.,
RA VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
RA VAUDIN M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL SCIENCE 265:2077-2082(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U00028; G458914; -.
DR PIR: S46675; S46675.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 571 AA; 65815 MW; 1A20600D CRC32;

Query Match 65.5%; Score 36; DB 1; Length 571;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WKAQKFLK 10
|| :|||
Db 198 WKTEKFLK 206

RESULT 12
YSX4_CAEL
ID YSX4_CAEL STANDARD; PRT; 434 AA.
AC Q10023;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 50.4 KD PROTEIN T28D9.4 IN CHROMOSOME II.
GN T28D9.4.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RA JOHNSTON R.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE CYSTEINE SYNTHASE/CYSTATHIONINE BETA-
CC SYNTHASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; U28738; G861265; -.
DR WORMPEP: T28D9.4; CE02069.
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
DOMAIN 151 154 POLY-LEU.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 263 283 POTENTIAL.
FT TRANSMEM 314 334 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
SQ SEQUENCE 434 AA; 50393 MW; 95639B3F CRC32;

Query Match 65.5%; Score 36; DB 1; Length 434;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KWKAKRFL 9
|| :|||
Db 3 RWKKPKRFL 11

RESULT 13
CBS_RAT
ID CBS_RAT STANDARD; PRT; 560 AA.
AC P32232;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 13-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CYSTATHIONINE BETA-SYNTHASE (EC 4.2.1.22) (SERINE SULFHYDRASE)
DE (BETA-THIONASE) (HEMOPROTEIN H-450).
GN CBS.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-LIVER:
RX MEDLINE; 92283859.
RA SWAROOP M., BRADLEY K., OHURA T., TAHARA T., ROBER M.D.,
RA ROSENBERG L.E., KRAUS J.P.;
RT "Rat cystathionine beta-synthase. Gene organization and alternative
RT splicing."
RL J. BIOL. CHEM. 267:11455-11461(1992).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-22 AND 38-47.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-LIVER;
RX MEDLINE; 91210211.
RA ISHIHARA S., MOROHASHI K.-I., SADANO H., KAWABATA S.-I., GOTOH O.,
RA OMURA T.;
RT "Molecular cloning and sequence analysis of cDNA coding for rat liver
RT hemoprotein H-450."
RL J. BIOCHEM. 108:899-902(1990).
CC -!- CATALYTIC ACTIVITY: L-SERINE + L-HOMOCYSTEINE -> CYSTATHIONINE +
CC H(2)O.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: FIRST STEP IN HOMOCYSTEINE TRANSULFURATION.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, AND BRAIN.
CC -!- PTM: BINDS COVALENTLY TO A HEME GROUP THROUGH A THIOLATE
CC LIGAND.
CC -!- ALTERNATIVE PRODUCTS: FOUR VARIANTS CAN BE FORMED BY ALTERNATIVE
CC SPLICING (TYPE I THROUGH IV). THE SEQUENCE OF TYPE I IS SHOWN.
CC -!- SIMILARITY: BELONGS TO THE CYSTEINE SYNTHASE/CYSTATHIONINE BETA-
CC SYNTHASE FAMILY.

```

CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 33-PHE-|-LEU-34 BOND IN THE  
CC ALPHA-CHAIN OF HEMOGLOBIN, LEADING TO DENATURATION OF MOLECULE.  
CC -!- SUBCELLULAR LOCATION: VACUOLAR. COULD BE FIRST ANCHORED TO THE  
CC MEMBRANE THROUGH ITS PROPEPTIDE BEFORE BEING RELEASED.  
CC -!- DEVELOPMENTAL STAGE: ERYTHROCYTIC STAGES  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE  
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X75787; G482941; -  
DR PIR: PT0434; PT0434.  
DR PROSITE: PS00141; ASP\_PROTEASE; 2.  
DR PFAM: PF00026; asp; 1.  
DR HSSP: P46925; LSME.  
KW HYDROLASE; ASPARTYL PROTEASE; GLYCOPROTEIN; ZYMOGEN; SIGNAL.  
FT SIGNAL 1 2  
FT PROPEP ? 124  
FT CHAIN 125 452 PLASMEPSIN 1.  
FT ACT\_SITE 157 157 BY SIMILARITY.  
FT ACT\_SITE 337 337 BY SIMILARITY.  
FT DISULFID 170 175 BY SIMILARITY.  
FT DISULFID 372 408 BY SIMILARITY.  
FT CARBOHYD 184 184 POTENTIAL.  
FT CARBOHYD 218 218 POTENTIAL.  
FT CARBOHYD 326 326 POTENTIAL.  
FT CARBOHYD 440 440 POTENTIAL.  
SQ SEQUENCE 452 AA; 51461 MW; AD192ADE CRC32;  
  
Query Match 65.5%; Score 36; DB 1; Length 452;  
Best Local Similarity 85.7%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 WKAKRF 8  
| | | | |  
DB 32 WKIQRF 38  
  
RESULT 9  
RK28\_ARATH  
ID RK28\_ARATH STANDARD; PRT; 140 AA.  
AC Q22795;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE 50S RIBOSOMAL PROTEIN L28, CHLOROPLAST PRECURSOR.  
GN RPL28 OR F4P9.22.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA.  
RA ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,  
RA SYKES S.M., MASON T.M., KERLAVAGE A.R., ADAMS M.D., SOMERVILLE C.R.,  
RA VENTER J.C.  
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AC002332; G2459427; -  
DR PFAM: PF00830; Ribosomal\_L28; 1.  
KW RIBOSOMAL PROTEIN; TRANSIT PEPTIDE; CHLOROPLAST.  
FT TRANSIT 1 63 CHLOROPLAST (BY SIMILARITY).  
FT CHAIN 64 140 50S RIBOSOMAL PROTEIN L28.  
SQ SEQUENCE 140 AA; 15699 MW; B600C960 CRC32;  
  
Query Match 65.5%; Score 36; DB 1; Length 140;  
Best Local Similarity 66.7%; Pred. No. 4.6;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 WKAKRF 10  
| | | | |  
DB 102 WEAGKRFV 110  
  
RESULT 10  
VEXB\_SALTI  
ID VEXB\_SALTI STANDARD; PRT; 264 AA.  
AC P43109.  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE VI POLYSACCHARIDE EXPORT INNER-MEMBRANE PROTEIN VEXB.  
GN VEXB.  
OS SALMONELLA TYPHI.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
OC SALMONELLA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GIFU 10007;  
RX MEDLINE: 93322324.  
RA HASHIMOTO I., LI N., YOKOYAMA H., EZAKI T.;  
RT "Complete nucleotide sequence and molecular characterization of Viab  
RT region encoding Vi antigen in Salmonella typhi.";  
RL J. BACTERIOL. 175:4456-4465(1993).  
CC -!- FUNCTION: MAY FORM AN ATP-DRIVEN VI POLYSACCHARIDE EXPORT  
CC APPARATUS. IN ASSOCIATION WITH THE VEXA, VEXC AND VEXD PROTEINS.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
CC (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE ABC-2 SUBFAMILY OF INTEGRAL MEMBRANE  
CC PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D14136; D1003706; -  
DR PROSITE: PS00890; ABC2\_MEMBRANE; 1.  
DR PFAM: PF01061; ABC2\_membrane; 1.  
KW POLYSACCHARIDE TRANSPORT; TRANSPORT; TRANSMEMBRANE; INNER MEMBRANE.  
FT TRANSMEM 38 58 POTENTIAL.  
FT TRANSMEM 65 85 POTENTIAL.  
FT TRANSMEM 120 140 POTENTIAL.  
FT TRANSMEM 149 169 POTENTIAL.  
FT TRANSMEM 191 211 POTENTIAL.  
FT TRANSMEM 239 259 POTENTIAL.  
SQ SEQUENCE 264 AA; 30429 MW; 52811E4C CRC32;  
  
Query Match 65.5%; Score 36; DB 1; Length 264;  
Best Local Similarity 66.7%; Pred. No. 8.8;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KWAKRF 9  
| | | | |



Query Match 74.5%; Score 41; DB 1; Length 481;  
 Best Local Similarity 70.0%; Pred. No. 1.9;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10  
 ||| :| |||  
 DB 115 KWVKRSFLK 124

RESULT 4  
 BPI\_RABIT STANDARD; PRT; 445 AA.  
 ID BPI\_RABIT AC Q28739;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE BACTERICIDAL PERMEABILITY-INCREASING PROTEIN (BPI) (FRAGMENT).  
 GN BPI.  
 OS ORYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NEW ZEALAND WHITE; TISSUE-BONE MARROW;  
 RA WEISS J., WEINRAUCH Y., LEVY O., FLYNN S.;  
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES  
 CC OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A  
 CC STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE  
 CC NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE  
 CC GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR  
 CC LEUKOCYTES (PMN) GRANULES (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES (BY  
 CC SIMILARITY).  
 CC -1- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF  
 CC THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE  
 CC MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE  
 CC RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-  
 CC AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED  
 CC RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE  
 CC INGESTED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; U61270; G1418284; -;  
 DR PROSITE; PS00400; LBP\_BPI\_CETP; PARTIAL.  
 DR PFAM; PF01273; LBP\_BPI\_CETP; 1.  
 DR HSP; P17213; LBPI  
 KW ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN.  
 FT SITE 198 203 CLEAVAGE SITES FOR ELASTASE (POTENTIAL).  
 FT TRANSHEM 327 347 POTENTIAL.  
 FT CARBOHYD 352 352 POTENTIAL.  
 SQ SEQUENCE 445 AA; 48837 MW; 5492B04C CRC32;  
 Query Match 70.9%; Score 39; DB 1; Length 445;  
 Best Local Similarity 60.0%; Pred. No. 4.2;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10  
 :|||:| :| :|  
 DB 80 RWKARKGFIK 89

RESULT 5  
 YDAK\_CAEL

ID YDAK\_CAEL STANDARD; PRT; 493 AA.  
 AC P90771; Q02645;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE PUTATIVE CYTOCHROME P450 C34B7.3 IN CHROMOSOME I (EC 1.14.-.-).  
 GN C34B7.3.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA HARRIS B.;  
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP REVISIONS.  
 RC STRAIN-BRISTOL N2;  
 RA JONES S.J.M.;  
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
 CC MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED  
 CC COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z83220; E1344512; -;  
 DR WORMPEP; C34B7.3; CE08567.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR PFAM; PF00087; P450; 1.  
 KW HYPOTHETICAL PROTEIN; OXIDOREDUCTASE; MONOOXYGENASE; HEME;  
 KW TRANSMEMBRANE.  
 FT TRANSHEM 1 21 POTENTIAL.  
 FT TRANSHEM 60 80 POTENTIAL.  
 FT TRANSHEM 290 310 POTENTIAL.  
 FT BINDING 440 440 HEME (BY SIMILARITY).  
 SQ SEQUENCE 493 AA; 56981 MW; D52DC3DD CRC32;  
 Query Match 69.1%; Score 38; DB 1; Length 493;  
 Best Local Similarity 85.7%; Pred. No. 7.2;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAKRKF 8  
 ||||:| :| :|  
 DB 118 WKAKRKF 124

RESULT 6  
 WCL\_NEUCR STANDARD; PRT; 1154 AA.  
 ID WCL\_NEUCR AC Q01371;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE WHITE COLLAR 1 PROTEIN (WCL).  
 GN WC-1.  
 OS NEUROSPORA CRASSA.  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;  
 OC SORDARIALES; SORDARIACEAE; NEUROSPORA.  
 RN [1]

```

CC -----
DR EMBL: J04739; G179529; ALT_INIT.
DR PIR: A29464; A29464.
DR PIR: A30909; A30909.
DR PIR: A33850; A33850.
DR PDB: 1BPI; 04-SEP-97.
DR MIM: 109195; -.
DR PROSITE: PS00400; LBP_BPI_CETP; 1.
DR PFAM: PF01273; LBP_BPI_CETP; 1.
KW ANTIBIOTIC; SIGNAL; TRANSMEMBRANE; GLYCOPROTEIN; 3D-STRUCTURE.
FT SIGNAL 1 27
FT CHAIN 28 483
FT SITE 236 241 BACTERICIDAL PERMEABILITY-INCREASING
FT CARBOHYD 365 385 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
FT SEQUENCE 483 AA; 53396 MW; 0AACEF65 CRC32;
SQ

Query Match 100.0%; Score 55; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKRFLK 10
Db 117 KWKAKRFLK 126

RESULT 2
BPI_BOVIN
ID BPI_BOVIN STANDARD; PRT; 482 AA.
AC P17453;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BACTERICIDAL PERMEABILITY-INCREASING PROTEIN PRECURSOR (BPI).
GN BPI.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDEA; BOVINA; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RX MEDLINE; 90272418.
RA LEONG S.R.; CAMERATO T.;
RT "Nucleotide sequence of the bovine bactericidal permeability
increasing protein (BPI).";
RL NUCLEIC ACIDS RES. 18:3052-3052(1990).
CC -1- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES
OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A
STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE
NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE
GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR
LEUKOCYTES (PMN) GRANULES (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES (BY
SIMILARITY).
CC -1- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF
THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE
MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE
RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-
AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED
RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE
INGESTED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: X52563; G139; -.
DR PIR: S10180; S10180.
DR PROSITE: PS00400; LBP_BPI_CETP; 1.
DR PFAM: PF01273; LBP_BPI_CETP; 1.
DR HSSP: P17213; LBPI.
KW ANTIBIOTIC; SIGNAL; MEMBRANE; GLYCOPROTEIN.
FT SIGNAL 1 26
FT CHAIN 27 482
FT SITE 235 240 BACTERICIDAL PERMEABILITY-INCREASING
FT CARBOHYD 62 62 POTENTIAL.
FT CARBOHYD 303 303 POTENTIAL.
FT CARBOHYD 375 375 POTENTIAL.
FT CARBOHYD 389 389 POTENTIAL.
FT CARBOHYD 463 463 POTENTIAL.
FT SEQUENCE 482 AA; 53432 MW; 26BA9CFA CRC32;
SQ

Query Match 80.0%; Score 44; DB 1; Length 482;
Best Local Similarity 70.0%; Pred. No. 0.53;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAKRFLK 10
Db 116 KWKAKRFLK 125

RESULT 3
LBP_MOUSE
ID LBP_MOUSE STANDARD; PRT; 481 AA.
AC Q61805;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).
GN LBP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE; 97289150.
RA LENGACHER S.; JONGENEEL C.V.; LE ROY D.; LEE J.D.; KRAVCHENKO V.;
RA ULEVITCH R.J.; GLAUSER M.P.; HEUMANN D.;
RT "Reactivity of murine and human recombinant LPS-binding protein (LBP)
within LPS and gram negative bacteria.";
RL J. INFLAMM. 47:165-172(1995).
CC -1- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL
LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER
MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS
TO INTERACT WITH THE CD14 RECEPTOR.
CC -1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99347; E253895; -.
DR MGD; MGI:1098776; LBP.
DR PROSITE: PS00400; LBP_BPI_CETP; 1.
DR PFAM: PF01273; LBP_BPI_CETP; 1.
DR HSSP: P17213; LBPI.
KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 24
FT CHAIN 25 481 LIPOPOLYSACCHARIDE-BINDING PROTEIN.
FT CARBOHYD 300 300 POTENTIAL.
FT CARBOHYD 355 355 POTENTIAL.
FT SEQUENCE 481 AA; 53312 MW; 973CFF58 CRC32;
SQ

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:07 ; Search time 71.09 Seconds  
(without alignments)  
3.976 Million cell updates/sec

Title: US-09-124-280A-27

Perfect score: 55

Sequence: 1 KWKAQRFLK 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	483	1 BPI_HUMAN	P17213 homo sapien
2	44	80.0	482	1 BPI_BOVIN	P17453 bos taurus
3	41	74.5	481	1 LBP_MOUSE	O61805 mus musculus
4	39	70.9	445	1 BPI_RABIT	O28739 oryctolagus
5	38	69.1	493	1 YDAK_CAEEL	P90771 caenorhabdi
6	37	67.3	1154	1 WCL_NEUCR	O01371 neurospora
7	36	65.5	538	1 CP18_DROME	O95078 drosophila
8	36	65.5	452	1 PLMI_PLAFA	P39898 plasmodium
9	36	65.5	140	1 RK28_ARATH	O22795 arabidopsis
10	36	65.5	264	1 VEXB_SALTI	P43109 salmonella
11	36	65.5	571	1 YHY4_YEAST	P38871 saccharomyc
12	36	65.5	434	1 YSX4_CAEEL	Q10023 caenorhabdi
13	35	63.6	560	1 CBS_RAT	P32232 rattus norv
14	35	63.6	1851	1 CCAT_DROME	P91645 drosophila
15	35	63.6	481	1 LBP_RAT	O63313 rattus norv
16	35	63.6	1002	1 ODOL_HUMAN	O02218 homo sapien
17	35	63.6	2344	1 POLN_RHDV	P27410 rabbit hemo
18	34	61.8	198	1 ATPE_MYCGA	P33256 mycoplasma
19	34	61.8	550	1 CBS_HUMAN	P35520 homo sapien
20	34	61.8	898	1 DPOL_BPT4	P04415 bacterioph
21	34	61.8	481	1 LBP_HUMAN	P18428 homo sapien
22	34	61.8	285	1 YDH5_SCHPO	O92350 schizosacch
23	34	61.8	461	1 YUL2_CAEEL	Q19802 caenorhabdi
24	33	60.0	883	1 CAPP_ECOLI	P00864 escherichia
25	33	60.0	286	1 CBPX_PEA	Q41005 plisum sativ
26	33	60.0	491	1 CPB1_RAT	P00176 rattus norv
27	33	60.0	491	1 CPB2_RAT	P04167 rattus norv
28	33	60.0	500	1 CPJ1_RABIT	P52786 oryctolagus
29	33	60.0	502	1 CPJ2_HUMAN	P51589 homo sapien
30	33	60.0	502	1 CPJ2_RAT	P51590 rattus norv
31	33	60.0	501	1 CPJ6_MOUSE	O54750 mus musculus
32	33	60.0	1895	1 GLS2_YEAST	P40989 saccharomyc
33	33	60.0	4451	1 GRSB_YACBR	P14688 bacillus br
34	33	60.0	2273	1 HFAL_YEAST	P32874 saccharomyc
35	33	60.0	467	1 IRF6_HUMAN	O14896 homo sapien
36	33	60.0	482	1 LBP_RABIT	P17454 oryctolagus
37	33	60.0	151	1 RK28_TOBAC	P30956 nicotiana t
38	33	60.0	164	1 RPO7_VACCV	P20984 vaccinia vi
39	33	60.0	164	1 RPO7_VARV	P33813 variola vir
40	33	60.0	287	1 Y011_MYCGE	P47257 mycoplasma
41	33	60.0	238	1 Y554_AQUAE	O66829 aquifex aeo
42	33	60.0	67	1 YPHE_BACSU	P50744 bacillus su
43	32	58.2	105	1 2SS2_CAPMA	P30233 capparid ma

ALIGNMENTS

RESULT 1

ID	BPI_HUMAN	STANDARD;	PRT;	483 AA.
AC	P17213;			
DT	01-AUG-1990 (REL. 15, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
DE	BACTERICIDAL PERMEABILITY-INCREASING PROTEIN PRECURSOR (BPI) (CAP 57).			
GN	BPI.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 28-64.			
RX	MEDLINE: 89255455.			
RA	GRAY P.W., FLAGS G., LEONG S.R., GUMINA R.J., WEISS J., OOI C.E.,			
RA	ELSBACH P.;			
RT	"Cloning of the cDNA of a human neutrophil bactericidal protein.			
RT	structural and functional correlations.";			
RL	J. BIOL. CHEM. 264:9505-9509(1989).			
RN	[2]			
RP	SEQUENCE OF 28-42.			
RX	MEDLINE: 88033057.			
RA	OOI C.E., WEISS J., ELSBACH P., FRANGIONE B., MANNION B.;			
RT	"A 25-kDa NH2-terminal fragment carries all the antibacterial			
RT	activities of the human neutrophil 60-kDa			
RL	bactericidal/permeability-increasing protein.";			
RL	J. BIOL. CHEM. 263:14891-14894(1987).			
RN	[3]			
RP	SEQUENCE OF 28-47.			
RX	MEDLINE: 89315847.			
RA	GABAY J.E., SCOTT R.W., CAMPANELLI D., GRIFFITH J., WILDE C.,			
RA	MARRA M.N., SEGER M., NATHAN C.F.;			
RT	"Antibiotic proteins of human polymorphonuclear leukocytes.";			
RT	PROC. NATL. ACAD. SCI. U.S.A. 86:5610-5614(1989).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).			
RX	MEDLINE: 97334442.			
RA	BEAUMER L.J., CARROLL S.F., EISENBERG D.;			
RT	"Crystal structure of human BPI and two bound phospholipids at 2.4-A			
RT	resolution.";			
RL	SCIENCE 276:1861-1864(1997).			
CC	-!- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES			
CC	OF GRAM-NEGATIVE BACTERIA: THIS SPECIFICITY MAY BE EXPLAINED BY A			
CC	STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE			
CC	NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE			
CC	GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.			
CC	-!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR			
CC	LEUKOCYTES (PMN) GRANULES.			
CC	-!- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES.			
CC	-!- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF			
CC	THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE			
CC	MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE			
CC	RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-			
CC	AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED			
CC	RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE			
CC	INGESTED.			
CC	-!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			

Q16181 homo sapien  
Q09571 caenorhabdi

**THIS PAGE BLANK (USPTO)**

C:Keywords: alternative splicing; carbon-oxygen lyase; hydro-lyase  
F:417-465/Domain: CBS homology <CBS2>

Query Match 63.6%; Score 35; DB 2; Length 561;  
Best Local Similarity 60.0%; Pred. No. 43;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KWKAKRFLK 10  
|| || ||  
Db 386 KMWLQKGFMK 395

Search completed: September 7, 1999, 23:22:42  
Job time: 949 sec



A:Residues: 1-2344 <MEY>  
 A:Cross-references: GB:M67473  
 C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
 C:Keywords: nucleotidyltransferase

Query Match 63.6%; Score 35; DB 1; Length 2344;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10  
 : : : |||||

Db 109 RWRRSRFLK 118

## RESULT 12

A38234

oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2) precursor - human  
 N:Alternate names: 2-oxoglutarate:lipoamide 2-oxidoreductase; alpha-ketoglutarate dehydrogenase  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 13-Nov-1998  
 C:Accession: A38234  
 R:Koike, K.; Urata, Y.; Goto, S.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 1963-1967, 1992  
 A:Title: Cloning and nucleotide sequence of the cDNA encoding human 2-oxoglutarate dehydrogenase  
 A:Reference number: A38234; MUID:92179301

A:Accession: A38234

A:Molecule type: mRNA

A:Residues: 1-1003 &lt;ROI&gt;

A:Cross-references: GB:D10523

A:Experimental source: fetal liver

A:Note: sequence extracted from NCBI backbone (NCBIP:87352)

C:Superfamily: oxoglutarate dehydrogenase (lipoamide); thiamine pyrophosphate-binding domain  
 C:Keywords: mitochondrion; oxidoreductase; thiamin pyrophosphate; tricarboxylic acid cycle  
 F:1-40/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
 F:41-1003/Product: oxoglutarate dehydrogenase (lipoamide) #status predicted <NAT>  
 F:408-453/Domain: thiamine pyrophosphate-binding domain homology <TPB>

Query Match 63.6%; Score 35; DB 2; Length 1003;  
 Best Local Similarity 62.5%; Pred. No. 79;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRF 8

: : : |||||

Db 271 KWSSEKRF 278

## RESULT 13

S55399

genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)  
 C:Species: rabbit hemorrhagic disease virus

A:Variety: isolate BS89

C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 08-Sep-1997

C:Accession: S55399

R:Rossi, C.

submitted to the EMBL Data Library, May 1995

A:Reference number: S55399

A:Accession: S55399

A:Molecule type: genomic RNA

A:Residues: 1-2344 &lt;ROS&gt;

A:Cross-references: EMBL:X87607; NID:g854640; PID:g854641

A:Experimental source: isolate BS89

C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C:Keywords: polyprotein

Query Match 63.6%; Score 35; DB 2; Length 2344;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10

: : : |||||

Db 109 RWRRSRFLK 118

## RESULT 14

S64740

genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)  
 N:Contains: VP60 protein  
 C:Species: rabbit hemorrhagic disease virus  
 A:Variety: isolate AST/89  
 C:Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 17-Mar-1999  
 C:Accession: S64740; S46944; S49018; S65012  
 R:Casals, R.; Martin-Alonso, J.; Boga, J.; Parra, F.  
 submitted to the EMBL Data Library, May 1995

A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by  
 A:Reference number: S64740  
 A:Accession: S64740

A:Molecule type: genomic RNA

A:Residues: 1-2344 &lt;CAS&gt;

A:Cross-references: EMBL:249271; NID:gl182032; PID:ei58171; PID:gl150552  
 A:Experimental source: isolate AST/89

R:Boga, J.; Casals, R.; Martin, M.; Martin-Alonso, J.; Carmones, R.; Prieto, M.; Parra  
 submitted to the EMBL Data Library, July 1993

A:Description: Molecular cloning, sequence and expression of the capsid protein gene  
 A:Reference number: S46944  
 A:Accession: S46944

A:Molecule type: genomic RNA

A:Residues: 1650-2344 &lt;BOG&gt;

A:Cross-references: EMBL:224757; NID:g515622; PID:g515623

A:Experimental source: isolate AST/89

R:Parra, F.; Boga, J.A.; Martin, M.S.; Casals, R.

Virus Res. 27, 219-228, 1993

A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus su  
 A:Reference number: S49018; MUID:93255896  
 A:Accession: S49018

A:Molecule type: genomic RNA

A:Residues: 1650-1796 &lt;PAR&gt;

A:Cross-references: EMBL:224757

A:Experimental source: isolate AST/89

A:Accession: S65012

A:Molecule type: protein

A:Residues: 1767-1779;1875-1877, 'X',1879-1881;1936-1938, 'X',1940-1941 <PAW>  
 C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
 C:Keywords: blocked amino end; polyprotein

Query Match 63.6%; Score 35; DB 2; Length 2344;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10

: : : |||||

Db 109 RWRRSRFLK 118

## RESULT 15

A42790

cystathionine beta-synthase (EC 4.2.1.22), splice form I - rat

N:Alternate names: beta-thionase; methylcysteine synthase; serine sulfhydrase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 15-Jan-1999

C:Accession: A42790

R:Swaroop, M.; Bradley, K.; Ohura, T.; Tahara, T.; Roper, M.D.; Rosenberg, L.E.; Krau

J. Biol. Chem. 267, 11455-11461, 1992

A:Title: Rat cystathionine beta-synthase. Gene organization and alternative splicing.

A:Reference number: A42790; MUID:92283859

A:Accession: A42790

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-561 &lt;SWA&gt;

A:Cross-references: GB:M88344; NID:gl364273; PID:g206597

A:Experimental source: strain Sprague-Dawley; liver

A:Note: sequence extracted from NCBI backbone (NCBIP:104763)

C:Superfamily: cystathionine beta-synthase; CBS homology

Query Match 65.5%; Score 36; DB 2; Length 354;  
Best Local Similarity 55.6%; Pred. No. 17;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRF 9  
: || || || ||  
Db 291 RWRMQRKYL 299

RESULT 7  
G36892  
Vi polysaccharide capsule transporter VexB - Salmonella typhi  
C:Species: Salmonella typhi  
C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 09-Sep-1997  
C:Accession: G36892  
R:Hashimoto, Y.; Li, N.; Yokoyama, H.; Ezaki, T.  
J. Bacteriol. 175, 4456-4465, 1993  
A:Title: Complete nucleotide sequence and molecular characterization of Viab region encoding Vi polysaccharide capsule transporter VexB  
A:Reference number: A36892; MUID:93322324  
A:Accession: G36892  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <HAS>  
A:Cross-references: GB:D14156; NID:g426443; PID:d1003706; PID:g426450  
C:Genetics:  
A:Gene: vexB

Query Match 65.5%; Score 36; DB 2; Length 264;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRF 9  
: || || || || ||  
Db 142 RWEAQRPF 150

RESULT 8  
F69326  
hypothetical protein AF0614 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
C:Accession: F69326  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
J. Bacteriol. 179, 137-141, 1997  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
A:Reference number: A69250; MUID:98049343  
A:Accession: F69326  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-93 <KLE>  
A:Cross-references: GB:AE001062; GB:AE000782; NID:g2689385; PID:g2650010; TIGR:AF0614

Query Match 65.5%; Score 36; DB 2; Length 93;  
Best Local Similarity 50.0%; Pred. No. 4.2;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAQKRF 10  
: || || || || ||  
Db 75 EWRMQRKYL 84

RESULT 9  
S46675  
hypothetical protein YHR184w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein H9186.1

Query Match 65.5%; Score 36; DB 2; Length 452;  
Best Local Similarity 85.7%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKQAQRFL 8  
: || || || || ||  
Db 32 WKIQKRF 38

RESULT 11  
RRWRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C:Species: rabbit hemorrhagic disease virus  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C:Accession: A41039  
R:Neyers, G.; Wirblich, C.; Thiel, H.J.  
Virology 184, 664-676, 1991  
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
A:Reference number: A41039; MUID:91361557  
A:Accession: A41039  
A:Molecule type: genomic RNA

Query Match 65.5%; Score 36; DB 2; Length 571;  
Best Local Similarity 55.6%; Pred. No. 28;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKQAQRFL 10  
: || || || || ||  
Db 198 WKTEKFLK 206

RESULT 10  
S41717  
aspartic hemoglobinase (EC 3.4.23.-) - Plasmodium falciparum  
C:Species: Plasmodium falciparum  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Jan-1999  
C:Accession: S41717; PT0434  
R:Francis, S.E.; Gluzman, I.Y.; Oksman, A.; Knickerbocker, A.; Mueller, R.; Bryant, M.  
EMBO J. 13, 306-317, 1994  
A:Title: Molecular characterization and inhibition of a Plasmodium falciparum aspartic  
A:Reference number: S41717; MUID:94147975  
A:Accession: S41717  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-452 <FRA>  
A:Cross-references: GB:X75787; NID:g482940; PID:g482941  
R:Goldberg, D.E.; Slater, A.F.G.; Beavis, R.; Chait, B.; Cerami, A.; Henderson, G.B.  
J. Exp. Med. 173, 961-969, 1991  
A:Title: Hemoglobin degradation in the human malaria pathogen Plasmodium falciparum:  
A:Reference number: PT0434; MUID:91178457  
A:Accession: PT0434  
A:Molecule type: protein  
A:Residues: 125-146 <GOL>  
A:Note: this sequence reveals 9 of 22 residues identical to the most specific mammal  
C:Comment: This enzyme plays a key role in the hemoglobinolytic pathway.  
C:Keywords: aspartic proteinase; hydrolase

Query Match 65.5%; Score 36; DB 2; Length 452;  
Best Local Similarity 85.7%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKQAQRFL 8  
: || || || || ||  
Db 32 WKIQKRF 38

RESULT 11  
RRWRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C:Species: rabbit hemorrhagic disease virus  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C:Accession: A41039  
R:Neyers, G.; Wirblich, C.; Thiel, H.J.  
Virology 184, 664-676, 1991  
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
A:Reference number: A41039; MUID:91361557  
A:Accession: A41039  
A:Molecule type: genomic RNA

C:Species: Saccharomyces cerevisiae  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 12-Dec-1997  
C:Accession: S46675  
R:Macri, C.  
submitted to the EMBL Data Library, February 1994  
A:Description: The sequence of S. cerevisiae cosmid 9186.  
A:Reference number: S46671  
A:Accession: S46675  
A:Molecule type: DNA  
A:Residues: 1-571 <MAC>  
A:Cross-references: EMBL:U00028; NID:g458909; PID:g458914; MIPS:YHR184w  
C:Genetics:  
A:Map position: 8R

Query Match 65.5%; Score 36; DB 2; Length 571;  
Best Local Similarity 55.6%; Pred. No. 28;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKQAQRFL 10  
: || || || || ||  
Db 198 WKTEKFLK 206

RESULT 10  
S41717  
aspartic hemoglobinase (EC 3.4.23.-) - Plasmodium falciparum  
C:Species: Plasmodium falciparum  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Jan-1999  
C:Accession: S41717; PT0434  
R:Francis, S.E.; Gluzman, I.Y.; Oksman, A.; Knickerbocker, A.; Mueller, R.; Bryant, M.  
EMBO J. 13, 306-317, 1994  
A:Title: Molecular characterization and inhibition of a Plasmodium falciparum aspartic  
A:Reference number: S41717; MUID:94147975  
A:Accession: S41717  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-452 <FRA>  
A:Cross-references: GB:X75787; NID:g482940; PID:g482941  
R:Goldberg, D.E.; Slater, A.F.G.; Beavis, R.; Chait, B.; Cerami, A.; Henderson, G.B.  
J. Exp. Med. 173, 961-969, 1991  
A:Title: Hemoglobin degradation in the human malaria pathogen Plasmodium falciparum:  
A:Reference number: PT0434; MUID:91178457  
A:Accession: PT0434  
A:Molecule type: protein  
A:Residues: 125-146 <GOL>  
A:Note: this sequence reveals 9 of 22 residues identical to the most specific mammal  
C:Comment: This enzyme plays a key role in the hemoglobinolytic pathway.  
C:Keywords: aspartic proteinase; hydrolase

Query Match 65.5%; Score 36; DB 2; Length 452;  
Best Local Similarity 85.7%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKQAQRFL 8  
: || || || || ||  
Db 32 WKIQKRF 38

RESULT 11  
RRWRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C:Species: rabbit hemorrhagic disease virus  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C:Accession: A41039  
R:Neyers, G.; Wirblich, C.; Thiel, H.J.  
Virology 184, 664-676, 1991  
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
A:Reference number: A41039; MUID:91361557  
A:Accession: A41039  
A:Molecule type: genomic RNA

Query Match 100.0%; Score 55; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.0062;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10  
| | | | | | | | | |  
Db 121 KWKAQKRFK 130

RESULT 2

S43383  
bactericidal/permeability-increasing protein - synthetic  
C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli  
C:Date: 20-Oct-1994 #sequence\_revision 15-Feb-1996 #text\_change 15-Feb-1996  
R:Qi, S.Y.; Li, Y.; O'Connor, C.D.  
Biochem. J. 298, 711-718, 1994  
A:Title: The region around residue 115 of human bactericidal/permeability-increasing protein  
of a gene coding for the active domain and characterization of recombinant proteins.  
A:Reference number: S43383  
A:Accession: S43383  
A:Molecule type: DNA  
A:Residues: 1-250 <QIS>

Query Match 100.0%; Score 55; DB 4; Length 250;  
Best Local Similarity 100.0%; Pred. No. 0.0031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10  
| | | | | | | | | |  
Db 91 KWKAQKRFK 100

RESULT 3

S10180  
bactericidal permeability-increasing protein precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 08-Sep-1997  
R:Leong, S.R.; Camerato, T.  
Nucleic Acids Res. 18, 3052, 1990  
A:Title: Nucleotide sequence of the bovine bactericidal permeability increasing protein  
A:Reference number: S10180; MUID:90272418  
A:Accession: S10180  
A:Molecule type: mRNA  
A:Residues: 1-482 <LEO>  
A:Cross-references: EMBL:X52563; NID:g138; PID:g139  
C:Superfamily: lipopolysaccharide-binding protein  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-482/Product: bactericidal permeability-increasing protein #status predicted <MAT>

Query Match 80.0%; Score 44; DB 2; Length 482;  
Best Local Similarity 70.0%; Pred. No. 0.73;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10  
| | | | | | | | | |  
Db 116 KWKAQKRFK 125

RESULT 4

S69206  
regulator protein white collar 1 - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 10-Oct-1997  
R:Ballario, P.; Vittorioso, P.; Magrelli, A.; Talora, C.; Cabibbo, A.; Macino, G.  
EMBO J. 15, 1650-1657, 1996  
A:Title: White collar-1, a central regulator of blue light responses in Neurospora, is a  
A:Reference number: S69206; MUID:96203083

A:Accession: S69206  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1154 <BAL>  
A:Cross-references: EMBL:X94300; NID:g1279576; PID:e256944; PID:g1480115  
C:Genetics:  
A:Introns: 967/3  
C:Superfamily: GATA-type zinc finger homology  
F:932-991/Domain: GATA-type zinc finger homology <GZF>

Query Match 67.3%; Score 37; DB 2; Length 1154;  
Best Local Similarity 85.7%; Pred. No. 38;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 WKAKKRF 8  
| | | | | | | |  
Db 1138 WKAKKRF 1144

RESULT 5

S52863  
DNA-binding protein R kappa B - human  
C:Species: Homo sapiens (man)  
C:Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 10-Sep-1997  
R:Nieters, A.; Bouwmeester, T.; Scheidereit, C.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: S52863  
A:Accession: S52863  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1324 <NIE>  
A:Cross-references: EMBL:X80878; NID:g695578; PID:g695579

Query Match 67.3%; Score 37; DB 2; Length 1324;  
Best Local Similarity 50.0%; Pred. No. 44;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10  
| | | | | | | | | |  
Db 198 EWRTQQRFLK 207

RESULT 6

S67642  
Probable arsenical pump-driving ATPase (EC 3.6.1.-) YDL100c - yeast (Saccharomyces ce  
N:Alternate names: hypothetical protein D2371  
C:Species: Saccharomyces cerevisiae  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 18-Sep-1998  
C:Accession: S67642; S67419  
R:Ballesta, J.P.G.; Remacha, M.; Soler-Mira, A.; Jimenez, A.; Garcia-Cantalejo, J.M.;  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67629  
A:Accession: S67642  
A:Molecule type: DNA  
A:Residues: 1-354 <BAL>  
A:Cross-references: EMBL:Z74148; NID:g1431137; PID:e253031; PID:g1431138; MIPS:YDL100  
A:Experimental source: strain S288C  
R:Boskovic, J.; Saiz, J.E.; Soler-Mira, A.; Garcia-Cantalejo, J.; Revuelta, J.L.; Jim  
submitted to the EMBL Data Library, February 1996  
A:Reference number: S67406  
A:Accession: S67419  
A:Molecule type: DNA  
A:Residues: 1-354 <BOS>  
A:Cross-references: EMBL:X95644; NID:g1199535; PID:e223235; PID:g1199549  
C:Genetics:  
A:Map position: 4L  
A:Note: YDL100c  
C:Superfamily: arsenical pump-driving ATPase  
C:Keywords: ATP; hydrolase; P-loop  
F:25-32/Region: nucleotide-binding motif A (P-loop)

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:41 ; Search time 105.14 Seconds  
(without alignments)  
3.811 Million cell updates/sec

Title: US-09-124-280A-27

Perfect score: 55

Sequence: 1 KWKAQRFLK 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database :  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	55	100.0	487	2	A30909	bactericidal/perme
2	55	100.0	250	4	S43383	bactericidal/perme
3	44	80.0	482	2	S10180	bactericidal/perme
4	37	67.3	1154	2	S69206	regulator protein
5	37	67.3	1324	2	S52863	DNA-binding protei
6	36	65.5	354	2	S67642	probable arsenical
7	36	65.5	264	2	G36892	Vi polysaccharide
8	36	65.5	33	2	F69326	hypothetical prote
9	36	65.5	571	2	S46575	hypothetical prote
10	36	65.5	452	2	S41717	aspartic hemoglobi
11	35	63.6	2344	1	RRWRH	genome polyprotein
12	35	63.6	1003	2	A38234	oxoglutarate dehyd
13	35	63.6	2344	2	S55399	genome polyprotein
14	35	63.6	2344	2	S64740	genome polyprotein
15	35	63.6	561	2	A42790	cystathionine beta
16	35	63.6	478	2	C42790	cystathionine beta
17	35	63.6	546	2	B42790	cystathionine beta
18	35	63.6	356	2	D42790	cystathionine beta
19	35	63.6	547	2	JX0145	hemoprotein H-450
20	35	63.6	481	2	I56246	lipopolysaccharide
21	35	63.6	373	2	E70458	translation releas
22	35	63.6	568	2	G02753	testis specific ba
23	35	63.6	1080	2	PC4394	DNA-directed DNA p
24	34	61.8	898	1	DJBPT4	cytochrome P450 -
25	34	61.8	487	2	C71417	cystathionine beta
26	34	61.8	552	2	A55760	lipopolysaccharide
27	34	61.8	481	2	A54136	lipopolysaccharide
28	34	61.8	477	2	A35843	H+-transporting AT
29	34	61.8	198	2	S24335	hypothetical prote
30	34	61.8	369	2	D70140	hypothetical prote
31	34	61.8	229	2	A64624	hypothetical prote
32	34	61.8	136	2	S76420	UTP-glucose glucos
33	34	61.8	462	2	T01732	hypothetical prote
34	34	61.8	131	2	D70448	hypothetical prote
35	33	60.0	491	1	O4RTPB	cytochrome P450 2B
36	33	60.0	451	1	O4RTP2	cytochrome P450 2B
37	33	60.0	164	1	RNV219	DNA-directed RNA p
38	33	60.0	883	1	QYEC	phosphoenolpyruvat
39	33	60.0	4452	1	YGBSG2	gramicidin S synth

ribosomal protein  
cytochrome P450 1b  
cytochrome P450 2B  
A5R protein - vari  
serine-type carbox  
probable transport

ALIGNMENTS

RESULT 1  
A30909  
bactericidal/permeability-increasing protein precursor - human  
N:Alternate names: 55K bactericidal protein  
C:Species: Homo sapiens (man)  
C>Date: 18-Apr-1989 #sequence\_revision 18-Apr-1989 #text\_change 17-Mar-1999  
C:Accession: A33850; B54136; A29464; A43600; A49716; A30909  
R:Gray, P.W.; Flagg, G.; Leong, S.R.; Gumina, R.J.; Weiss, J.; Ooi, C.E.; Elsbach, P.  
J. Biol. Chem. 264, 9505-9509, 1989  
A:Title: Cloning of the cDNA of a human neutrophil bactericidal protein. Structural a  
A:Reference number: A33850; MUID:89255455  
A:Accession: A33850  
A:Molecule type: mRNA  
A:Residues: 1-487 <GRA>  
A:Cross-references: GB:J04739; NID:q179528; PID:q179529  
R:Wilde, C.G.; Seilhamer, J.J.; McGrogan, M.; Ashton, N.; Snable, J.L.; Lane, J.C.; L  
J. Biol. Chem. 269, 17411-17416, 1994  
A:Title: Bactericidal/permeability-increasing protein and lipopolysaccharide (LPS)-bi  
A:Reference number: A54136; MUID:94292492  
A:Accession: B54136  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-374, L', 376-487 <WIL>  
A:Experimental source: HL-60 cells  
A:Note: sequence extracted from NCBI backbone (NCBIP:149855)  
R:Ooi, C.E.; Weiss, J.; Elsbach, P.; Frangione, B.; Mannion, B.  
J. Biol. Chem. 262, 14891-14894, 1987  
A:Title: A 25-kDa amino-terminal fragment carries all the antibacterial activities of  
A:Reference number: A29464; MUID:88033057  
A:Accession: A29464  
A:Molecule type: protein  
A:Residues: 32-51 <OOI>  
A:Experimental source: neutrophils  
R:Wasilik, K.R.; Skubitz, K.M.; Gray, B.H.  
Infect. Immun. 59, 4193-4200, 1991  
A:Title: Comparison of granule proteins from human polymorphonuclear leukocytes which  
A:Reference number: A43600; MUID:92040097  
A:Accession: A43600  
A:Molecule type: protein  
A:Residues: 32-52, 'R' <WAS>  
R:Little, R.G.; Keiner, D.N.; Lim, E.; Burke, D.J.; Conlon, P.J.  
J. Biol. Chem. 269, 1865-1872, 1994  
A:Title: Functional domains of recombinant bactericidal/permeability increasing prote  
A:Reference number: A49716; MUID:94124531  
A:Accession: A49716  
A:Molecule type: protein  
A:Residues: 32-130/132-141; 143-165; 202-215, 'E', 217-225 <LIT>  
C:Comment: The bactericidal/permeability-increasing protein (BPI) is a 60 kD membrane  
which is specific for gram-negative bacteria. BPI has a high affinity for lipopolysac  
between BPI and an LPS-binding protein from liver and cholesterol ester transfer prot  
C:Genetics:  
A:Gene: GDB:BPI  
A:Cross-references: GDB:131572; OMIM:109195  
A:Map position: 20q11.23-20q12  
C:Superfamily: lipopolysaccharide-binding protein  
C:Keywords: antibacterial; cytotoxin; glycoprotein; heparin binding; neutrophil  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-487/Product: bactericidal permeability-increasing protein #status predicted <MAT  
F:32-51/Region: bactericidal #status predicted  
F:380/Binding site: carbohydrate (Asn) #status predicted

**THIS PAGE BLANK (USPTO)**

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,611A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/273,401  
FILING DATE: 11-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/125,651  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 32251  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "BPI.17"  
US-08-311-611A-18

Query Match 100.0%; Score 55; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10  
Db 6 KWKAQKRFK 15

RESULT 15  
US-08-311-611A-19  
Sequence 19, Application US/08311611A  
Patent No. 5523288  
GENERAL INFORMATION:  
APPLICANT: Cohen, Jonathan  
APPLICANT: Kung, Ada H.C.  
APPLICANT: Lambert, Jr., Lewis H.  
TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
TITLE OF INVENTION: Infection by Administration of  
TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 227  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,611A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/273,401  
FILING DATE: 11-JUL-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/125,651  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 32251  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "BPI.18"  
US-08-311-611A-19

Query Match 100.0%; Score 55; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10  
Db 6 KWKAQKRFK 15

Search completed: September 7, 1999, 23:07:46  
Job time: 1744 sec

```

; Sequence 16, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration Of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311.611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "BPI.15"
US-08-311-611A-16

Query Match 100.0%; Score 55; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10
Db 6 KWKAQKRFK 15

RESULT 13
US-08-311-611A-17
; Sequence 17, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration Of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311.611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "BPI.16"
US-08-311-611A-17

Query Match 100.0%; Score 55; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10
Db 6 KWKAQKRFK 15

RESULT 14
US-08-311-611A-18
; Sequence 18, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration Of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

;  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "BPI.58"  
US-08-311-611A-9

Query Match 100.0%; Score 55; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFLK 10  
Db 7 KWKAQKRFLK 16  
|||||

## RESULT 10

US-08-311-611A-10  
; Sequence 10, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "BPI.65 oxidized"  
US-08-311-611A-10

Query Match 100.0%; Score 55; DB 1; Length 17;

Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KWKAQKRFLK 10  
Db 7 KWKAQKRFLK 16  
|||||

## RESULT 11

US-08-311-611A-11  
; Sequence 11, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "BPI.3"  
US-08-311-611A-11

Query Match 100.0%; Score 55; DB 1; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFLK 10  
Db 18 KWKAQKRFLK 27  
|||||

RESULT 12  
US-08-311-611A-16



;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/311.611A  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/273.401  
;; FILING DATE: 11-JUL-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/125.651  
;; FILING DATE: 22-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sharp, Jeffrey S.  
;; REGISTRATION NUMBER: 31,879  
;; REFERENCE/DOCKET NUMBER: 32251  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/474-6300  
;; TELEFAX: 312/474-0448  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 15 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "BPI.2"  
US-08-311-611A-7

Query Match 100.0%; Score 55; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAKRFLK 10  
Db 6 KWKAKRFLK 15

RESULT 8  
US-08-311-611A-8  
; Sequence 8, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/permeability-Increasing  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311.611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273.401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125.651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:

;; NAME: Sharp, Jeffrey S.  
;; REGISTRATION NUMBER: 31,879  
;; REFERENCE/DOCKET NUMBER: 32251  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/474-6300  
;; TELEFAX: 312/474-0448  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "BPI.8"  
US-08-311-611A-8  
  
Query Match 100.0%; Score 55; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00091;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAKRFLK 10  
Db 1 KWKAKRFLK 10

RESULT 9  
US-08-311-611A-9  
; Sequence 9, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/permeability-Increasing  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311.611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273.401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125.651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids

QY 1 KWKAQRFLK 10  
| | | | |  
Db 121 KWKAQRFLK 130

RESULT 5  
US-08-414-924-2  
; Sequence 2, Application US/08414924  
; Patent No. 5494896  
; GENERAL INFORMATION:  
; APPLICANT: Hansbrough, John F.  
; TITLE OF INVENTION: Method of Treating Conditions  
; TITLE OF INVENTION: Associated With Burn Injuries  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/414,924  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 27129/32297  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-414-924-2

Query Match 100.0%; Score 55; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFLK 10  
| | | | |  
Db 121 KWKAQRFLK 130

RESULT 6  
US-08-311-611A-6  
; Sequence 6, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "Domain II"  
US-08-311-611A-6

Query Match 100.0%; Score 55; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFLK 10  
| | | | |  
Db 26 KWKAQRFLK 35

RESULT 7  
US-08-311-611A-7  
; Sequence 7, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```
; APPLICANT: Grinna, Lynn S
; TITLE OF INVENTION: Stable Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Products and Pharmaceutical
; TITLE OF INVENTION: Compositions Containing the Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/013.801
; FILING DATE: 02 FEB 1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 27129/30911
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/346-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-013-801-2

Query Match 100.0%; Score 55; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQRFLK 10
Db 121 KWKAQRFLK 130

RESULT 3
US-08-072-063-2
; Sequence 2, Application US/08072063
; Patent No. 5439807
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Grinna, Lynn S
; APPLICANT: Horwitz, Arnold
; TITLE OF INVENTION: BPI-Immunoglobulin Fusion Proteins
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072.063
; FILING DATE: 19930519
```

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: 30659
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-072-063-2

Query Match 100.0%; Score 55; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQRFLK 10
Db 121 KWKAQRFLK 130

RESULT 4
US-08-212-132-2
; Sequence 2, Application US/08212132
; Patent No. 5447913
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; APPLICANT: Ammons, William Steve
; TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Dimer Products
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212.132
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/31735
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-212-132-2

Query Match 100.0%; Score 55; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:46 ; Search time 82.37 Seconds  
(without alignments)  
1.198 Million cell updates/sec.

Title: US-09-124-280A-27

Perfect score: 55  
Sequence: 1 KWAQKRFK 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	487	1	US-08-030-644-2
2	55	100.0	487	1	US-08-013-801-2
3	55	100.0	487	1	US-08-072-063-2
4	55	100.0	487	1	US-08-212-132-2
5	55	100.0	487	1	US-08-414-924-2
6	55	100.0	35	1	US-08-311-611A-6
7	55	100.0	15	1	US-08-311-611A-7
8	55	100.0	10	1	US-08-311-611A-8
9	55	100.0	16	1	US-08-311-611A-9
10	55	100.0	17	1	US-08-311-611A-10
11	55	100.0	27	1	US-08-311-611A-11
12	55	100.0	15	1	US-08-311-611A-16
13	55	100.0	15	1	US-08-311-611A-17
14	55	100.0	15	1	US-08-311-611A-18
15	55	100.0	15	1	US-08-311-611A-19
16	55	100.0	15	1	US-08-311-611A-20
17	55	100.0	15	1	US-08-311-611A-49
18	55	100.0	15	1	US-08-311-611A-51
19	55	100.0	24	1	US-08-311-611A-52
20	55	100.0	29	1	US-08-311-611A-53
21	55	100.0	20	1	US-08-311-611A-54
22	55	100.0	25	1	US-08-311-611A-55
23	55	100.0	20	1	US-08-311-611A-57
24	55	100.0	20	1	US-08-311-611A-58
25	55	100.0	26	1	US-08-311-611A-65
26	55	100.0	17	1	US-08-311-611A-66
27	55	100.0	17	1	US-08-311-611A-68
28	55	100.0	487	1	US-08-311-611A-69
29	55	100.0	24	1	US-08-311-611A-70
30	55	100.0	487	1	US-08-311-611A-146
31	55	100.0	30	1	US-08-311-611A-149
32	55	100.0	20	1	US-08-311-611A-150
33	55	100.0	20	1	US-08-311-611A-151
34	55	100.0	14	1	US-08-311-611A-164
35	55	100.0	12	1	US-08-311-611A-227
36	55	100.0	487	1	US-08-173-968-2
37	55	100.0	221	1	US-08-173-968-3
38	55	100.0	487	1	US-08-232-527-2
39	55	100.0	35	1	US-08-372-783-6

40 55 100.0 15 1 US-08-372-783-7 Sequence 7, Appli  
41 55 100.0 10 1 US-08-372-783-8 Sequence 8, Appli  
42 55 100.0 16 1 US-08-372-783-9 Sequence 9, Appli  
43 55 100.0 17 1 US-08-372-783-10 Sequence 10, Appli  
44 55 100.0 27 1 US-08-372-783-11 Sequence 11, Appli  
45 55 100.0 487 3 PCT-US95-09262-206 Sequence 206, App

## ALIGNMENTS

## RESULT 1

US-08-030-644-2

; Sequence 2, Application US/08030644

; Patent No. 5348942

; GENERAL INFORMATION:

; APPLICANT: Little, Roger G. II

; APPLICANT: Gazzano-Santoro, Helene

; APPLICANT: Parent, James Brian

; TITLE OF INVENTION: Therapeutic uses of

; TITLE OF INVENTION: Bactericidal/Permeability-Increasing Protein Products

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/030.644

; FILING DATE: 19930312

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Sharp, Jeffrey S.

; REGISTRATION NUMBER: 31,879

; REFERENCE/DOCKET NUMBER: 31229

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300

; TELEFAX: 312/474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 487 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-030-644-2

Query Match 100.0%; Score 55; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.031;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KWAQKRFK 10

Db 121 KWAQKRFK 130

## RESULT 2

US-08-013-801-2

; Sequence 2, Application US/08013801

; Patent No. 5420019

; GENERAL INFORMATION:

; APPLICANT: Theofan, Georgia

; APPLICANT: Horwitz, Arnold

; APPLICANT: Burke, David

; APPLICANT: Baltaian, Manik

**THIS PAGE BLANK (USPTO)**

Search completed: September 7, 1999, 22:49:57  
Job time: 7878 sec

---

CC rheumatoid arthritis, and for treating gram-negative bacterial infection.  
 CC The peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipopolysaccharide in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipopolysaccharide or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 20 AA;

Query Match 100.0%; Score 55; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0035;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KWKAKRFLK 10  
 | | | | | | | | | |  
 Db 1 KWKAKRFLK 10

## RESULT 14

ID R63727 standard; peptide; 15 AA.

AC R63727;  
 DT 03-MAY-1995 (first entry)  
 DE BPI derived peptide, BPI.9.  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW contraception; malignant; tumour cell; inflammatory disease; T-cell;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipopolysaccharide; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 PN W09420532-A.  
 PD 15-SEP-1994.

PF 11-MAR-1994; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.  
 PA (XOMA ) XOMA CORP.  
 PI Little RG;  
 DR WPI; 94-302964/37.

PT New human bactericidal permeability increasing peptides - derived  
 PT from the functional domains of BPI and having BPI activities such  
 PT as bactericidal activity

PS Claim 8; Page 158; 254pp; English.  
 CC The sequences given in R6382-750, R62087-100 and R62491-500 are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC protein (BPI). The sequences given in R63727-35 are derived from  
 CC positions 65-99 of BPI and may form dimers or trimers. Peptides  
 CC such as these may be used as heparin binding agents, for neutralising  
 CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.  
 CC associated with ocular retinopathy, for inhibiting endothelial cell  
 CC proliferation, for treating a chronic inflammatory disease state, eg.  
 CC rheumatoid arthritis, and for treating gram-negative bacterial infection.  
 CC The peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipopolysaccharide in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipopolysaccharide or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 15 AA;

Query Match 100.0%; Score 55; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0027;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KWKAKRFLK 10  
 | | | | | | | | | |  
 Db 6 KWKAKRFLK 15

## RESULT 15

ID R63730 standard; peptide; 20 AA.

AC R63730;  
 DT 03-MAY-1995 (first entry)  
 DE BPI derived peptide, BPI.46.  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW contraception; malignant; tumour cell; inflammatory disease; T-cell;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipopolysaccharide; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 PN W09420532-A.  
 PD 15-SEP-1994.

PF 11-MAR-1994; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.

PA (XOMA ) XOMA CORP.  
 PI Little RG;  
 DR WPI; 94-302964/37.

PT New human bactericidal permeability increasing peptides - derived  
 PT from the functional domains of BPI and having BPI activities such  
 PT as bactericidal activity  
 PS Claim 8; Page 159; 254pp; English.  
 CC The sequences given in R6382-750, R62087-100 and R62491-500 are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC protein (BPI). The sequences given in R63727-35 are derived from  
 CC positions 65-99 of BPI and may form dimers or trimers. Peptides  
 CC such as these may be used as heparin binding agents, for neutralising  
 CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.  
 CC associated with ocular retinopathy, for inhibiting endothelial cell  
 CC proliferation, for treating a chronic inflammatory disease state, eg.  
 CC rheumatoid arthritis, and for treating gram-negative bacterial infection.  
 CC The peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipopolysaccharide in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipopolysaccharide or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 20 AA;

Query Match 100.0%; Score 55; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0035;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KWKAKRFLK 10  
 | | | | | | | | | |  
 Db 11 KWKAKRFLK 20

## RESULT 11

R57901  
ID R57901 standard; Protein; 487 AA.  
AC R57901;  
DT 29-MAR-1995 (first entry)  
DE Bactericide BPI.  
KW Bactericidal-permeability-increasing protein; bactericide;  
KW fusion protein; Gram-negative bacterium; infection.  
OS Homo sapiens.  
FH Key  
FT peptide Location/Qualifiers  
FT 1..31  
FT /label= Sig\_peptide  
PN WO9418323-A.  
PD 18-AUG-1994.  
PF 02-FEB-1994. U01235.  
PR 02-FEB-1993; US-013801.  
PA (XOMA ) XOMA CORP.  
PI Baltalan M, Burke D, Grinna L, Horwitz A, Theofan G;  
DR WPI: 94-279744/34.  
DR N-PSDB; Q67270.  
PT Bactericidal-permeability-increasing protein analogs with  
PT Cys132 or Cys135 replaced - also fusion proteins and C-terminally  
PT truncated forms, have increased stability and are useful to treat  
PT gram-negative bacterial infection  
PS Disclosure; Page 52-54; 77pp; English.  
CC Analogs, fusion proteins and C-terminal truncated forms of human BPI  
CC are claimed that show improved resistance to dimerization and  
CC adduct formation.  
SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.061;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFELK 10

DB 121 KWKAQKRFELK 130

## RESULT 12

R62370  
ID R62370 standard; Protein; 487 AA.  
AC R62370;  
DT 27-APR-1995 (first entry)  
DE Recombinant bactericidal/permeability-increasing protein (rBPI-50).  
KW N-terminal; recombinant; Bactericidal/permeability protein; rBPI;  
KW signal protein; human; rBPI-23; binding assay; heparin-binding; domain;  
KW Limulus amoebocyte lysate; LAL; inhibition; bactericidal activity;  
KW chronic inflammatory disease; arthritis; psoriasis; Crohn's disease;  
KW inflammatory bowel disease; asthma.  
OS Homo sapiens.  
FH Key  
FT peptide Location/Qualifiers  
FT 1..31  
FT /note= "Signal peptide"  
FT protein 32..487  
FT /note= "mature protein"  
FT domain 17..55  
FT /note= "LAL inhibiting domain"  
FT binding\_site 21..55  
FT /note= "Heparin binding site"  
FT binding\_site 65..107  
FT /note= "Heparin binding site"  
FT domain 73..99  
FT /note= "LAL inhibiting domain"  
FT active\_site 85..99  
FT /note= "Bactericidally active fragment"  
FT domain 137..163  
FT /note= "LAL inhibiting domain"  
FT binding\_site 137..171  
FT /note= "Heparin binding site"  
PN WO9420128-A.  
PD 15-SEP-1994.

PF 11-MAR-1994; U02401.  
PR 12-MAR-1993; US-030644.  
PR 15-JUL-1993; US-093202.  
PA (XOMA ) XOMA CORP.  
PI Gazzano-Santoro H, Little RG, Parent JB;  
DR WPI: 94-302679/37.  
DR N-PSDB; Q73382.  
PT neutralise anticoagulant effect of heparin, to inhibit  
PT angiogenesis or cell proliferation, as contraceptive and to reduce  
PT inflammation  
PS Disclosure; Page 43-45; 78pp; English.  
CC This sequence represents the N-terminal of recombinant bactericidal/  
CC permeability (rBPI) protein. This expression product encoding the 31  
CC residue signal protein and the first 199 amino acids of the N-terminal  
CC of mature human BPI has a molecular weight of 23 kD and is designated  
CC rBPI-23. In binding assays rBPI-23 was shown to have three separate  
CC functional domains with heparin-binding activity, at least three major  
CC domains with significant Limulus amoebocyte lysate (LAL) inhibition and  
CC one region with bactericidal activity. BPI, or active fragments of it  
CC may be used to treat and prevent the effects of chronic inflammatory  
CC disease states such as arthritis, psoriasis, inflammatory bowel disease,  
CC Crohn's disease and asthma.  
SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.061;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFELK 10

DB 121 KWKAQKRFELK 130

## RESULT 13

R63731  
ID R63731 standard; peptide; 20 AA.  
AC R63731;  
DT 03-MAY-1995 (first entry)  
DE BPI derived peptide, BPI.47.  
KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
KW rheumatoid arthritis; malignant; tumour cell; inflammatory disease; T-cell;  
KW lipopolysaccharide; circulation; gram-negative bacteria; infection; cytokine;  
KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
OS Homo sapiens.  
PN WO9420532-A.  
PD 15-SEP-1994.  
PF 11-MAR-1994; U02465.  
PR 12-MAR-1993; US-030644.  
PR 15-JUL-1993; US-093202.  
PR 14-JAN-1994; US-183222.  
PA (XOMA ) XOMA CORP.  
PI Little RG;  
DR WPI: 94-302964/37.  
PT New human bactericidal permeability increasing peptides - derived  
PT from the functional domains of BPI and having BPI activities such  
PT as bactericidal activity  
PS Claim 8; Page 159; 254pp; English.  
CC The sequences given in R63682-750, R62087-100 and R62491-500 are  
CC peptides derived from human bactericidal/permeability-increasing  
CC protein (BPI). The sequences given in R63727-35 are derived from  
CC positions 65-99 of BPI and may form dimers or trimers. Peptides  
CC such as these may be used as heparin binding agents, for neutralising  
CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.  
CC associated with ocular retinopathy, for inhibiting endothelial cell  
CC proliferation, for contraception, for inhibiting malignant tumour cell  
CC proliferation, for treating a chronic inflammatory disease state, eg.



FT WO9323434-A.  
 PN 25-NOV-1993.  
 PD 19-MAY-1993; U04754.  
 PR 19-MAY-1992; US-885911.  
 PA (XOMA ) XOMA CORP.  
 PI Grinna LS, Horwitz A, Theofan G;  
 DR WPI: 93-386485/48.  
 DR N-PSDB: Q52488.  
 PT New fusion proteins for treating bacterial infections - comprising  
 a bactericidal/permeability-increasing protein and an immunoglobulin  
 heavy chain constant domain  
 PS Claim 20; Page 48-50; 75pp; English.  
 CC This sequence is encoded by the plasmid PING4512. PING4512 was  
 CC constructed using the primer sequence given in Q52480. Plasmid  
 CC PING4503 was cut with ALENI, the ends made blunt with T4 DNA  
 CC polymerase, and the DNA then cut with SalI. The resultant approx.  
 CC 700bp SalI/blunt fragment contains the 30 bp 5' untranslated region  
 CC and the DNA encoding the signal peptide and first 191 amino acids of  
 CC bactericidal/permeability-increasing protein (BPI). Igg heavy chain  
 CC (HC) sequences, which also included the hinge region of the Igg HC  
 CC were PCR amplified from plasmid PMB27 using these primers. Plasmids  
 CC such as pING4512 may be used to produce fusion proteins which are  
 CC useful for the treatment of gram negative bacterial infections and  
 CC their sequelae including endotoxin related shock and conditions  
 CC associated with it, such as disseminated intravascular coagulation,  
 CC anemia, thrombocytopenia, leukopenia, adult respiratory distress  
 CC syndrome, renal failure, hypotension, fever and metabolic acidosis.  
 CC Proving BPI or a fragment of it, as part of the fusion with an  
 CC immunoglobulin heavy chain constant region provides the potential  
 CC advantages of Fc receptor binding, bivalent binding to lipopoly-  
 CC saccharide complement binding and increasing placental transfer.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. NO. 0.061;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKRFLK 10  
 DB 121 KWKAKRFLK 130  
 RESULT 9  
 R43886  
 ID R43886 standard; Protein; 487 AA.  
 AC R43886;  
 DT 01-JUN-1994 (first entry)  
 DE Human bactericidal/permeability-increasing protein.  
 KW Bactericidal/permeability-increasing protein;  
 KW endotoxin binding protein.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT protein 32..487  
 FT /label= human BPI protein  
 FT misc\_difference 1..31  
 FT /label= signal peptide  
 PN WO9323540-A.  
 PD 25-NOV-1993.  
 PF 19-MAY-1993; U04752.  
 PR 19-MAY-1992; US-885501.  
 PA (XOMA ) XOMA CORP.  
 PI Grinna LS;  
 DR WPI: 93-386572/48.  
 DR P-PSDB: Q52268.  
 PT Recovering endotoxin binding protein from transformed cell  
 PT cultures - by retention on cation exchange material and elution,  
 PT giving simpler recovery and higher yields, for treating bacterial  
 PT infections  
 PS Disclosure; Page 27-28; 38pp; English.  
 CC Human bactericidal/permeability increasing protein (an endotoxin  
 CC binding protein) is useful for treatment of bacterial infections.

SQ Sequence 487 AA;  
 Query Match 100.0%; Score 55; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. NO. 0.061;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KWKAKRFLK 10  
 DB 121 KWKAKRFLK 130  
 RESULT 10  
 R63687  
 ID R63687 standard; peptide; 35 AA.  
 AC R63687;  
 DT 03-MAY-1995 (first entry)  
 DE BPI derived peptide, domain II.  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipopolysaccharide; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 PN WO9420532-A.  
 PD 15-SEP-1994.  
 PF 11-MAR-1994; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.  
 PA (XOMA ) XOMA CORP.  
 PI Little RG;  
 DR WPI: 94-302964/37.  
 PT New human bactericidal permeability increasing peptides - derived  
 PT from the functional domains of BPI and having BPI activities such  
 PT as bactericidal activity  
 PS Claim 6; Page 157; 254pp; English.  
 CC The sequences given in R63682-750, R62087-100 and R62491-500 are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC protein (BPI). The sequences given in R63687-726 are derived from  
 CC positions 65-99 of BPI. Peptides such as these may be used as  
 CC heparin binding agents, for neutralising the anti-coagulant effect  
 CC of heparin, for inhibiting angiogenesis, eg. associated with ocular  
 CC retinopathy, for inhibiting endothelial cell proliferation, for  
 CC contraception, for inhibiting malignant tumour cell proliferation,  
 CC for treating a chronic inflammatory disease state, eg. rheumatoid  
 CC arthritis, and for treating gram-negative bacterial infection. The  
 CC peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipopolysaccharide in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipopolysaccharide or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 SQ Sequence 35 AA;

Query Match 100.0%; Score 55; DB 1; Length 35;  
 Best Local Similarity 100.0%; Pred. NO. 0.0057;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KWKAKRFLK 10  
 DB 26 KWKAKRFLK 35

CC The hBPI holoprotein was isolated from human promyelocytic leukemia  
 CC cells (HL-60) and BPI cDNA was prep'd. The N-terminal fragments of  
 CC BPI exhibit all the anti-bacterial and membrane permeability-  
 CC increasing functions contained in the complete molecule (as well as  
 CC the LPS-neutralising function) but are substantially smaller. In  
 CC addition, the N-terminal fragment, as well as the C-terminal  
 CC fragment(s) are capable of inhibiting the prodn. and/or release of  
 CC tumour necrosis factor (TNF) triggered by LPS in whole blood. The  
 CC inventors claim the DNA sequence comprising bps 721-1491 of Q25593.  
 CC They also claim the following fragments of MATURE hBPI: 1-99 (I);  
 CC 1-203 (II); 200-456 (III); 204-456 (IV). NB: The above fragments  
 CC are labeled with respect to the first AA of the mature sequence,  
 CC which corresp. to residue 32 of R24777.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.061;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQRFLK 10  
 |||||

Db 121 KWKAQRFLK 130

RESULT 6

R33776 ID R33776 standard; protein; 368 AA.

AC R33776; 22-JUL-1993 (first entry)

DE BPI from p337.

KW Bacterial/permeability increasing protein; lipopolysaccharide; LPS;  
 KW inhibitor; endotoxin-related shock; disseminated intravascular  
 KW coagulation; anaemia; thrombocytopenia; ARDS; renal failure;  
 KW endotoxemia; adult respiratory distress syndrome.

OS Homo sapiens.

FH Key Location/Qualifiers

FT peptide 1..32

FT region /note= "signal peptide"

FT /note= "mature protein"

FT WO9305797-A.

PN 01-APR-1993.

PD 28-SEP-1992; U08234.

PF 27-SEP-1991; US-766566.

PR (INCY-) INCYTE PHARM INC.

PA Marra MN, Scott RW, Snable JL, Wilde CG;

PI WPI; 93-117241/14.

DR Compens. contg. a bactericidal and permeability increasing

PT protein - solubilised in a lipid carrier to maintain stably an

PT active form

PS Disclosure: Fig 6; 53pp; English.

CC The human bactericidal/permeability increasing protein may be used  
 CC in a compsn. solubilised in a lipid carrier. The BPI binds  
 CC lipopolysaccharides and inhibits the immunostimulatory and toxic  
 CC activities of LPS in vitro and in vivo. The lipid compsn. stably  
 CC maintains the BPI in a sol. active form. The compsn. can be used  
 CC for detecting or quantifying LPS in a sample, coating a surgical  
 CC tool, coating an implantable invasive device, decontaminating a  
 CC fluid contg. LPS and treating endotoxin related -shock,  
 CC -disseminated intravascular coagulation, -anaemia, -thrombocytopenia,  
 CC -adult respiratory stress syndrome or -renal failure, or for

CC preventing endotoxaemia.

CC See also R33777, R33729-31.

CC Sequence 368 AA;

Query Match 100.0%; Score 55; DB 1; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 0.047;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQRFLK 10  
 |||||

Db 121 KWKAQRFLK 130

RESULT 7

ID R33777 standard; protein; 243 AA.

AC R33777;

DT 22-JUL-1993 (first entry)

DE BPI from p212.

KW Bacterial/permeability increasing protein; lipopolysaccharide; LPS;

KW inhibitor; endotoxin-related shock; disseminated intravascular

KW coagulation; anaemia; thrombocytopenia; ARDS; renal failure;

KW endotoxemia; adult respiratory distress syndrome.

OS Homo sapiens.

FH Key Location/Qualifiers

FT peptide 1..32

FT region /note= "signal peptide"

FT /note= "mature protein"

FT WO9305797-A.

PN 01-APR-1993.

PD 28-SEP-1992; U08234.

PR 27-SEP-1991; US-766566.

PA (INCY-) INCYTE PHARM INC.

PI Marra MN, Scott RW, Snable JL, Wilde CG;

DR WPI; 93-117241/14.

PT Compens. contg. a bactericidal and permeability increasing

PT protein - solubilised in a lipid carrier to maintain stably an

PT active form

PS Disclosure: Fig 7; 53pp; English.

CC The human bactericidal/permeability increasing protein may be used

CC in a compsn. solubilised in a lipid carrier. The BPI binds

CC lipopolysaccharides and inhibits the immunostimulatory and toxic

CC activities of LPS in vitro and in vivo. The lipid compsn. stably

CC maintains the BPI in a sol. active form. The compsn. can be used

CC for detecting or quantifying LPS in a sample, coating a surgical

CC tool, coating an implantable invasive device, decontaminating a

CC fluid contg. LPS and treating endotoxin related -shock,

CC -disseminated intravascular coagulation, -anaemia, -thrombocytopenia,

CC -adult respiratory stress syndrome or -renal failure, or for

CC preventing endotoxaemia.

CC See also R33776, R33729-31.

SQ Sequence 243 AA;

Query Match 100.0%; Score 55; DB 1; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQRFLK 10  
 |||||

Db 121 KWKAQRFLK 130

RESULT 8

ID R44420

AC R44420 standard; Protein; 487 AA.

DT 03-JUN-1994 (first entry)

DE PING4512 encoded fusion protein.

KW Polymerase chain reaction; primer; amplify; PCR; plasmid; PING4512;

KW PING4503; bactericidal/permeability-increasing protein; BPI; IgG;

KW heavy; chain; PMB27; fusion protein; gram negative; renal failure;

KW bacterial infection; endotoxin related shock; metabolic acidosis;

KW disseminated intravascular coagulation; anaemia; thrombocytopenia;

KW leukopenia; adult respiratory distress syndrome; hypotension; fever;

KW constant region; lipopolysaccharide complement binding;

KW placental transfer; FC receptor binding.

OS Synthetic.

FH Key Location/Qualifiers

FT peptide 1..31

FT /note= "Signal peptide"

FT protein 32..487

FT region 1..31  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991; U05758.  
 PR 13-AUG-1990; US-567016.  
 PR 05-APR-1991; US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI: 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Disclosure: Fig 13; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxemia and endotoxic shock.  
 SQ Sequence 368 AA;

Query Match 100.0%; Score 55; DB 1; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 0.047;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KWKAQKRFLLK 10  
 DB 121 KWKAQKRFLLK 130

RESULT 3  
 R21843  
 ID R21843 standard; Protein; 396 AA.  
 AC R21843;  
 DT 22-JUL-1992 (first entry)  
 DE Sequence of CHO-BPI a biologically active variant of BPI, a human  
 DE 57kd protein which binds to the outer membrane of susceptible gram  
 DE negative bacteria (BPI).  
 DE Endotoxin binding protein; diagnosis; therapy; endotoxemia;  
 KW endotoxic shock.  
 OS Homo sapiens.  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991; U05758.  
 PR 13-AUG-1990; US-567016.  
 PR 05-APR-1991; US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI: 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Claim 5; Fig 24; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxemia and endotoxic shock.  
 SQ Sequence 396 AA;

Query Match 100.0%; Score 55; DB 1; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KWKAQKRFLLK 10  
 DB 90 KWKAQKRFLLK 99

RESULT 4  
 R21844  
 ID R21844 standard; Protein; 456 AA.  
 AC R21844;  
 DT 22-JUL-1992 (first entry)  
 DE Sequence of BPI (DP linkage), a biologically active variant of BPI, a  
 DE 57kd protein which binds to the outer membrane of susceptible gram  
 DE negative bacteria (BPI).  
 DE Endotoxin binding protein; diagnosis; therapy; endotoxemia;  
 KW endotoxic shock.  
 OS Homo sapiens.  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991; U05758.  
 PR 13-AUG-1990; US-567016.  
 PR 05-APR-1991; US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI: 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Claim 52; Fig 25; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxemia and endotoxic shock.  
 SQ Sequence 456 AA;

Query Match 100.0%; Score 55; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 0.058;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KWKAQKRFLLK 10  
 DB 90 KWKAQKRFLLK 99

RESULT 5  
 R24777  
 ID R24777 standard; Protein; 487 AA.  
 AC R24777;  
 DT 02-JAN-1992 (first entry)  
 DE Sequence of human bactericidal/permeability-increasing protein  
 DE (hBPI) derived from human promyelocytic leukemia cells (HL-60).  
 DE Bactericidal; endotoxin inhibitor; anti-microbial.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 PN WO9209621-A.  
 PD 11-JUN-1992.  
 PF 03-DEC-1991; U09033.  
 PR 03-DEC-1990; US-621473.  
 PR 26-AUG-1991; US-754204.  
 PA (UYNV) UNIV NEW YORK STATE.  
 PI Elsbach P, Weiss J;  
 DR WPI: 92-217016/26.  
 DR N-PSDB; 025593.  
 PT Antibacterial fragments of bactericidal-permeability-increasing  
 PT protein - for treating gram negative bacterial infections esp. in  
 PT immuno suppressed patients, with endotoxin activity  
 PS Claim 16; Fig 5; 63pp; English.

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:49:57 ; Search time 135.78 seconds  
(without alignments)  
1.744 Million cell updates/sec

Title: US-09-124-280A-27

Perfect score: 55

Sequence: 1 KWKAQKRFLK 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	487	1 P94636	Human bactericidal
2	55	100.0	368	1 R21840	Sequence of p337,
3	55	100.0	396	1 R21843	Sequence of CHO-BP
4	55	100.0	456	1 R21844	Sequence of BPI (D
5	55	100.0	487	1 R24777	Sequence of human
6	55	100.0	368	1 R33776	BPBP from p337..Co
7	55	100.0	243	1 R33777	BPBP from p212..Co
8	55	100.0	487	1 R44420	PING4512-encoded f
9	55	100.0	487	1 R43886	Human bactericidal
10	55	100.0	35	1 R63687	BPI derived peptid
11	55	100.0	487	1 R57901	Bactericidal BPI. B
12	55	100.0	487	1 R62370	Recombinant bacter
13	55	100.0	20	1 R63731	BPI derived peptid
14	55	100.0	15	1 R63727	BPI derived peptid
15	55	100.0	20	1 R63730	BPI derived peptid
16	55	100.0	24	1 R63729	BPI derived peptid
17	55	100.0	21	1 R63728	BPI derived peptid
18	55	100.0	26	1 R63735	BPI derived peptid
19	55	100.0	24	1 R62495	BPI derived peptid
20	55	100.0	29	1 R62496	BPI derived peptid
21	55	100.0	24	1 R62497	BPI derived peptid
22	55	100.0	15	1 R63715	BPI derived peptid
23	55	100.0	17	1 R63720	BPI derived peptid
24	55	100.0	10	1 R63689	BPI derived peptid
25	55	100.0	16	1 R63690	BPI derived peptid
26	55	100.0	17	1 R63691	BPI derived peptid
27	55	100.0	27	1 R63692	BPI derived peptid
28	55	100.0	15	1 R63693	BPI derived peptid
29	55	100.0	15	1 R63694	BPI derived peptid
30	55	100.0	15	1 R63695	BPI derived peptid
31	55	100.0	15	1 R63696	BPI derived peptid
32	55	100.0	15	1 R63697	BPI derived peptid
33	55	100.0	487	1 R62344	Recombinant 23 kD
34	55	100.0	15	1 R63688	BPI derived peptid
35	55	100.0	487	1 R67998	Holoprotein rBP150
36	55	100.0	487	1 R71290	Human bactericidal
37	55	100.0	368	1 R71291	Human bactericidal
38	55	100.0	243	1 R71292	Human bactericidal
39	55	100.0	10	1 R71798	Peptide neutralisi
40	55	100.0	199	1 R68825	BPI(1-159)/lipopol
41	55	100.0	197	1 R68929	Lipopolysaccharide
42	55	100.0	199	1 R68931	LBP(1-87)/BPI(88-1
43	55	100.0	198	1 R68933	BPI(1-147)/LBP(147

ALIGNMENTS

RESULT 1

P94636 standard; protein; 487 AA.

AC P94636;

DT 29-JUN-1990 (first entry)

DE Human bactericidal/permeability-increasing (BPI) holoprotein

KW Polypeptide fragment; bacteraemia; sepsis; gram-negative bacteria;

KW bactericidal/permeability-increasing (BPI) holoprotein;

KW mammalian myeloid blood cells; polymorphonuclear leukocytes (PMN);

KW human promyelocytic leukaemia cells (HL-60).

OS Homo Sapiens.

FH Key Location/Qualifiers

FT peptide 1..31

FT protein 32..487

FT region 153..155

FT /note="potential glycosylation site"

FT region 379..381

FT /note="potential glycosylation site"

PN W08901486-A.

PD 23-FEB-1989.

PF 09-AUG-1988; U02700.

PR 05-AUG-1988; US-228035, US-084335.

PA (UNYK) New York Univ.

PI Elsbach P, Weiss J;

DR WPI: 89-068849/09.

DR N-PSDB; N92766

PT Bacterial-permeability increasing holoprotein fragments -

PT used for treatment of diseases caused by gram negative bacteria

PT such as bacteraemia or sepsis.

PS Disclosure: ; 4lpp; English.

CC The purified, isolated cDNA which encodes this sequence was isolated

CC from human promyelocytic leukaemia cells (HL-60) and is also claimed.

CC Biologically active BPI protein fragments can be used for the

CC treatment of mammals suffering from diseases caused by gram-negative

CC bacteria such as bacteraemia or sepsis. They have good selectivity and

CC lack of cytotoxicity towards cells other than gram-negative bacteria.

CC The fragments may be used in a prepn. further comprising rifampicin,

CC penicillin or cephalosporin. The BPI holoproteins used in prodn. of

CC the fragments may be obtd. from mammalian myeloid blood cells, eg

CC polymorphonuclear leukocytes (PMN), or by recombinant DNA techniques.

CC BPI fragments are prepd. by incubating the holoprotein under protein

CC cleaving conditions or using recombinant DNA techniques.

SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.061;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFLK 10

|||||

Db 121 KWKAQKRFLK 130

RESULT 2

R21840

ID R21840 standard; protein; 368 AA.

AC R21840;

DT 22-JUL-1992 (first entry)

DE Sequence of p337, a biologically active fragment of the human

DE 57kd protein which binds to the outer membrane of susceptible gram

DE negative bacteria (BPI).

KW Endotoxin binding protein; diagnosis; therapy; endotoxemia;

KW endotoxin shock.

OS Homo sapiens.

FH Key Location/Qualifiers

**THIS PAGE BLANK (USPTO)**

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF045641; AAC02576.1; -  
 DR PFAM; PF00270; DEAD; 1.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 KW ATP-binding; RNA-binding; Helicase.  
 SQ SEQUENCE 646 AA; 74177 MW; 31753542 CRC32;

Query Match 68.0%; Score 34; DB 5; Length 646;  
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKR 7  
 ||:||||  
 Db 111 KWKAQKR 117

Search completed: September 7, 1999, 22:47:36  
 Job time: 7956 sec

DE HYPOTHETICAL 18.4 KD PROTEIN.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PCC6803;  
 RA TABATA S.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PCC6803;  
 RX MEDLINE; 97061201.  
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,  
 RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,  
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,  
 RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,  
 RA TABATA S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions."  
 RL DNA Res 3:109-136(1996).  
 DR EMBL; D90915; BAAL8549.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 156 AA; 18386 MW; 832B91B3 CRC32;

Query Match 68.0%; Score 34; DB 2; Length 156;  
 Best Local Similarity 66.7%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
 :||:||||  
 Db 52 RWEAQRKGL 60

RESULT 15  
 O45076 PRELIMINARY; PRT; 646 AA.  
 AC O45076;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE F53H1.1 PROTEIN.  
 GN F53H1.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., JIER M., JOHNSTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., LAISTER N., LATREILLE P.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LAURENCE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans."  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA BECK C., KRAMER J., KEPPLER D.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;

```

Db 262 WKAQKGF 268

RESULT 10
P72197 PRELIMINARY; PRT; 1723 AA.
AC P72197;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE LYS-GINGIPAIN.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Cytophagales; Bacteroidaceae; Porphyromonas.
RN [1]
RP SEQUENCE FROM N.A.
RA PAVLOFF N., PEMBERTON P.A., POTEMPA J., CHEN W.C.A., PIKE R.N.,
RA PROCHAZKA V., KIEFER M.C., TRAVIS J., BARR P.J.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U54691; AAA99810.1; -.
DR PFAM; PF01364; Peptidase_C25; 2.
SQ SEQUENCE 1723 AA; 186830 MW; A5CD4AD0 CRC32;

Query Match 68.0%; Score 34; DB 2; Length 1723;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WKAQKRF 8
|||||
Db 262 WKAQKGF 268

RESULT 11
O07442 PRELIMINARY; PRT; 1732 AA.
ID O07442;
AC O07442;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE LYSINE-SPECIFIC CYSTEINE PROTEINASE.
GN PRTK.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Cytophagales; Bacteroidaceae; Porphyromonas.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-W50;
RA SLAKESKI N., CLEAL S.M., REYNOLDS E.C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U75366; AAB60809.1; -.
DR PFAM; PF01364; Peptidase_C25; 2.
SQ SEQUENCE 1732 AA; 187913 MW; 7FB74472 CRC32;

Query Match 68.0%; Score 34; DB 2; Length 1732;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WKAQKRF 8
|||||
Db 262 WKAQKGF 268

RESULT 12
O51304 PRELIMINARY; PRT; 369 AA.
ID O51304;
AC O51304;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 42.8 KD PROTEIN.
GN BB0325.
OS Borrelia burgdorferi (Lyme disease spirochete).

```

---

```

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., OUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AE001139; AAC66711.1; -.
DR TIGR; BB0325; -.
KW Hypothetical protein.
SQ SEQUENCE 369 AA; 42782 MW; 26C485A1 CRC32;

Query Match 68.0%; Score 34; DB 2; Length 369;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 WKAQKREL 9
|:|::|:|
Db 324 WRAERRFI 331

RESULT 13
P72194 PRELIMINARY; PRT; 1723 AA.
ID P72194;
AC P72194;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE LYS-GINGIPAIN.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Cytophagales; Bacteroidaceae; Porphyromonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-381;
RX MEDLINE; 97044756.
RA OKAMOTO K., KADOWAKI T., NAKAYAMA K., YAMAMOTO K.;
RT "Cloning and sequencing of the gene encoding a novel lysine-specific
RT cysteine proteinase (Lys-gingipain) in Porphyromonas gingivalis:
RT structural relationship with the arginine-specific cysteine
RT proteinase (Arg-gingipain)."
RL J. Biochem. 120:398-406(1996).
DR EMBL; D83258; BAAL1870.1; -.
DR PFAM; PF01364; Peptidase_C25; 2.
SQ SEQUENCE 1723 AA; 187261 MW; B589E9D6 CRC32;

Query Match 68.0%; Score 34; DB 2; Length 1723;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WKAQKRF 8
|||||
Db 262 WKAQKGF 268

RESULT 14
P74448 PRELIMINARY; PRT; 156 AA.
ID P74448;
AC P74448;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)

```

Query Match 70.0%; Score 35; DB 10; Length 589;  
 Best Local Similarity 66.7%; Pred. No. 64;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAQKRF 9  
 |||||  
 DB 331 KWKSQHTFL 339

RESULT 6  
 ID O29641 PRELIMINARY; PRT; 93 AA.  
 AC O29641;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-AUG-1998 (Tremblrel. 07, Last annotation update)  
 DE HYPOTHETICAL 11.1 KD PROTEIN.  
 GN AF0614.

OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 OC Archaeoglobus.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE; 98049343.

RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,  
 RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,  
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
 RA KIRKNESS E., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
 RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLOBEK A., ZHOU L.,  
 RA OVERBECK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,  
 RA COTTON M.D., SPRIGGS T., ARTACH P., KAINE B.P., SYKES S.M.,  
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
 RA MASON T.N., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,  
 RA VENTER J.C.;

RT "The complete genome sequence of the hyperthermophilic,  
 RT sulphate-reducing archaeon Archaeoglobus fulgidus.";

RL Nature 390:364-370(1997).  
 DR EMBL; AE001062; AAB90629.1; -

DR TIGR; AF0614; -  
 KW Hypothetical protein.

SQ SEQUENCE 93 AA; 11148 MW; 0961504C CRC32;

Query Match 68.0%; Score 34; DB 1; Length 93;  
 Best Local Similarity 55.8%; Pred. No. 17;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAQKRF 9  
 |||||  
 DB 75 EWKSKRFI 83

RESULT 7  
 ID Q51817 PRELIMINARY; PRT; 1732 AA.  
 AC Q51817;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE PORPHYRAIN.  
 GN PRTP.

OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; Cytophagales; Bacteroidaceae; Porphyromonas.

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=W12.  
 RX MEDLINE; 96213011.

RA BARKOCY-CALLAGHER G.A., HAN N., PATTI J.M., WHITLOCK J.,  
 RA PROGULSKE-FOX A., LANTZ M.S.;

RT "Analysis of the prtp gene encoding porphyrin, a cysteine proteinase

RT of Porphyromonas gingivalis.";  
 RL J. Bacteriol. 178:2734-2741(1996).  
 DR EMBL; U42210; AAB06565.1; -  
 DR PFAM; PF01364; Peptidase\_C25; 2.  
 SQ SEQUENCE 1732 AA; 187874 MW; 0C0C402C CRC32;

Query Match 68.0%; Score 34; DB 2; Length 1732;  
 Best Local Similarity 85.7%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAKQRF 8  
 |||||  
 DB 262 WKAKQGF 268

RESULT 8  
 ID O52050 PRELIMINARY; PRT; 1732 AA.  
 AC O52050;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE LYSINE SPECIFIC CYSTEINE PROTEASE.  
 OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; Cytophagales; Bacteroidaceae; Porphyromonas.

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=W83;  
 RA LEWIS J.P., MACRINA F.L.;

RL Infect. Immun. 0:0-0(1997).  
 DR EMBL; AF017059; AAC26523.1; -

DR PFAM; PF01364; Peptidase\_C25; 2.  
 KW Protease.

SQ SEQUENCE 1732 AA; 187930 MW; 48A8A4F CRC32;

Query Match 68.0%; Score 34; DB 2; Length 1732;  
 Best Local Similarity 85.7%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAKQRF 8  
 |||||  
 DB 262 WKAKQGF 268

RESULT 9  
 ID P96966 PRELIMINARY; PRT; 364 AA.  
 AC P96966;  
 DT 01-MAY-1997 (Tremblrel. 03, Created)  
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)  
 DT 01-AUG-1998 (Tremblrel. 07, Last annotation update)  
 DE CYSTEINE PROTEASE.  
 GN KGP-381.

OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; Cytophagales; Bacteroidaceae; Porphyromonas.

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=381;

RA HAN N., LEPIE G., WHITLOCK J., WOJCIECHOWSKI L., PROGULSKE-FOX A.;

RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U68468; AAB49690.1; -

KW Protease.

SQ SEQUENCE 364 AA; 40160 MW; 429CAF36.CRC32;

Query Match 68.0%; Score 34; DB 2; Length 364;  
 Best Local Similarity 85.7%; Pred. No. 62;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAKQRF 8  
 |||||



```

RN SEQUENCE FROM N.A.
RP STRAIN=VF5;
RX MEDLINE: 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUWAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUWAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE000758; AAC07656.1; -.
DR PFAM; PF00472; RF-1; 1.
SQ SEQUENCE 373 AA; 43309 MW; 921E417E CRC32;

Query Match 70.0%; Score 35; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWAQAK 6
Db 361 KWAQAK 366
|||||

RESULT 3
Q92538 PRELIMINARY; PRT; 1691 AA.
AC Q92538.
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE MYELOBLAST KIAA0248 (FRAGMENT).
GN KIAA0248.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RX MEDLINE: 97191544.
RA NAGASE T., SEKI N., ISHIKAWA K., OHIRA M., KAWARABAYASI Y., OHARA O.,
RA TANAKA A., KOTANI H., MIYAJIMA N., NOMURA N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT analysis of cDNA clones from cell line KG-1 and brain."
RL DNA Res. 3:321-329(1996).
DR EMBL; D87435; BAA13379.1; -.
DR PFAM; PF01369; Sec7; 1.
FT NON_TER 1
SQ SEQUENCE 1691 AA; 187806 MW; 8D8EC432 CRC32;

Query Match 70.0%; Score 35; DB 4; Length 1691;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWAQAK 7
Db 58 KWKQKR 64
|||||

RESULT 4
Q27465 PRELIMINARY; PRT; 548 AA.
ID Q27465.
AC Q27465.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE SIMILAR TO CYTOCHROME P450 (EC 1.14.14.1).
GN B0304.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA GRISEL C.; (NOV-1995) to the EMBL/GenBank/DBJ databases.
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; U39472; AAA80134.1; -.
DR PFAM; PF00067; P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.
FT BINDING 471 471 HEME (BY SIMILARITY).
SQ SEQUENCE 548 AA; 63665 MW; C4F8339E CRC32;

Query Match 70.0%; Score 35; DB 5; Length 548;
Best Local Similarity 62.5%; Pred. No. 59;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWAQAKRF 8
Db 146 KWREQRRF 153
||:|:|:|

RESULT 5
Q9ZRT1 PRELIMINARY; PRT; 589 AA.
ID Q9ZRT1.
AC Q9ZRT1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE GAMMA RESPONSE 1 PROTEIN.
GN GR 1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA DEVEAUX Y., KAZMAIER M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ131708; CAA10484.1; -.
SQ SEQUENCE 589 AA; 67472 MW; 798F6B88 CRC32;

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:34 ; Search time 160.2 Seconds  
(without alignments)  
3.458 Million cell updates/sec

Title: US-09-124-280A-26

Perfect score: 50

Sequence: 1 KWKAKRFL 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- SPTREMBL\_10:\*
- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phase:\*
- 10: sp.plant:\*
- 11: sp.podent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	36	72.0	354	3	Q12154	Q12154 saccharomyc
2	35	70.0	373	2	O67695	O67695 aquifex ao
3	35	70.0	1691	4	Q92538	Q92538 homo sapien
4	35	70.0	548	5	Q27465	Q27465 caenorhabdi
5	35	70.0	589	10	Q928T1	Q928T1 arabidopsi
6	34	68.0	93	1	O29641	O29641 archaeoglob
7	34	68.0	1732	2	Q51817	Q51817 porphyromon
8	34	68.0	1732	2	O52050	O52050 porphyromon
9	34	68.0	364	2	P96966	P96966 porphyromon
10	34	68.0	1723	2	P72197	P72197 porphyromon
11	34	68.0	1732	2	O07442	O07442 porphyromon
12	34	68.0	369	2	O51304	O51304 borrelia bu
13	34	68.0	1723	2	P72194	P72194 porphyromon
14	34	68.0	156	2	P74448	P74448 synecocyst
15	34	68.0	646	5	O45076	O45076 caenorhabdi
16	34	68.0	1038	5	O61199	O61199 caenorhabdi
17	34	68.0	660	5	Q17248	Q17248 boophilus m
18	34	68.0	680	5	Q94196	Q94196 caenorhabdi
19	34	68.0	74	7	O97983	O97983 trionyx sin
20	33	66.0	141	1	Q50529	Q50529 methanobact
21	33	66.0	337	2	O67278	O67278 aquifex ao
22	33	66.0	2273	3	O42823	O42823 saccharomyc
23	33	66.0	469	3	O59710	O59710 schizosacch
24	33	66.0	1704	4	Q92473	Q92473 homo sapien
25	33	66.0	1704	4	Q99758	Q99758 homo sapien
26	33	66.0	248	5	O18360	O18360 drosophila
27	33	66.0	248	5	O18361	O18361 drosophila
28	33	66.0	248	5	O18362	O18362 drosophila
29	33	66.0	245	5	O18418	O18418 drosophila

30	33	66.0	247	5	O18419	O18419 drosophila
31	33	66.0	248	5	O18667	O18667 drosophila
32	33	66.0	245	5	O18668	O18668 drosophila
33	33	66.0	248	5	O18675	O18675 drosophila
34	33	66.0	337	5	O16046	O16046 drosophila
35	33	66.0	322	5	O16047	O16047 drosophila
36	33	66.0	690	5	O61142	O61142 plasmodium
37	33	66.0	499	5	O45219	O45219 caenorhabdi
38	33	66.0	462	10	O04622	O04622 arabidopsi
39	33	66.0	1093	11	O70535	O70535 rattus norv
40	33	66.0	467	11	P97431	P97431 mus musculu
41	33	66.0	395	11	O64579	O64579 rattus norv
42	33	66.0	575	11	O88482	O88482 mus musculu
43	33	66.0	491	11	O64584	O64584 rattus norv
44	33	66.0	164	12	O89610	O89610 variola vir
45	33	66.0	459	13	O57578	O57578 xenopus lae

ALIGNMENTS

RESULT 1  
Q12154  
ID Q12154 PRELIMINARY: PRT: 354 AA.  
AC  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-NOV-1996 (TREMREL. 01, Last annotation update)  
DE ORF YDL100C FROM LEFT ARM OF CHROMOSOME IV.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomycetes.  
RP [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-FY1679;  
RA BOSKOVIC J., SAIZ J.E., SOLER-MIRA A., GARCIA-CANTALEJO J.,  
RA REVUELTA J.L., JIMENEZ A., BALESTA J.P.G., DEL REY F., REMACHA M.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RN SEQUENCE FROM N.A.  
RA BALLESTA J.P.G., REMACHA M., SOLER-MIRA A., JIMENEZ A.,  
RA GARCIA-CANTALEJO J.M., BOSKOVIC J., DEL REY F., REVUELTA J.L.,  
RA BUITRAGO M.J., SANZ J.E.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RA MIPS;  
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X95644; CAA64913.1; -;  
DR EMBL; 274148; CAA98667.1; -;  
SQ SEQUENCE 354 AA; 39353 MW; 63AB1475 CRC32;

Query Match 72.0%; Score 36; DB 3; Length 354;  
Best Local Similarity 55.6%; Pred. No. 26;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAKRFL 9  
:|:|:|:|:  
Db 291 RWMQKKYL 299

RESULT 2  
O67695  
ID O67695 PRELIMINARY: PRT: 373 AA.  
AC  
DT 01-AUG-1998 (TREMREL. 07, Created)  
DT 01-AUG-1998 (TREMREL. 07, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE PEPTIDE CHAIN RELEASE FACTOR RF-2.  
GN PRFB.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificales; Aquificaceae; Aquifex.

FT CONFLICT 89 89 A -> V (IN G15382).  
SQ SEQUENCE 898 AA: 103609 MW: 8C05B769 CRC32;

Query Match 68.0%; Score 34; DB 1; Length 898;  
Best Local Similarity 71.4%; Pred. No. 50;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 WKAKQRF 8  
      |||:|:  
Db 699 WKAKKRY 705

RESULT 15  
YDH5\_SCHPO  
ID YDH5\_SCHPO STANDARD; PRT; 285 AA.  
AC 092350;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 32.8 KD PROTEIN C6G9.05 IN CHROMOSOME I.  
GN SPAC6G9.05.  
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;  
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;  
OC SCHIZOSACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., CONNOR R.E.;  
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: BELONGS TO THE UPF0035 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; 281317; E276540; -.  
DR PROSITE; PS01293; UPF0035; 1.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 285 AA: 32772 MW: 50807CD0 CRC32;

Query Match 68.0%; Score 34; DB 1; Length 285;  
Best Local Similarity 55.6%; Pred. No. 15;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KWAKQRF 9  
      |||:|:  
Db 6 RWKGHOREL 14

Search completed: September 7, 1999, 23:59:07  
Job time: 519 sec

DR PIR: S24335; S24335.  
 DR PFAM: PF00430; ATP-synt\_B; 1.  
 KW HYDROGEN ION TRANSPORT; TRANSMEMBRANE; CF(0).  
 SQ SEQUENCE 198 AA; 22723 MW; B38FA3E7 CRC32;

Query Match 68.0%; Score 34; DB 1; Length 198;

Best Local Similarity 75.0%; Pred. No. 10;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 WKQAQKREL 9

II IIII  
 DB 67 WKTKREL 74

RESULT 13

BPI\_RABIT  
 ID BPI\_RABIT STANDARD; PRT; 445 AA.

AC Q28739; (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE BACTERICIDAL PERMEABILITY-INCREASING PROTEIN (BPI) (FRAGMENT).  
 GN BPI.  
 OS ORYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-NEW ZEALAND WHITE; TISSUE-BONE MARROW;  
 RA WEISS J., WEINRAUCH Y., LEVI O., FLYNN S.;  
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES  
 OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A

CC STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE  
 CC NEGATIVELY CHARGED LIPOLYSACCHARIDES THAT ARE UNIQUE TO THE

CC GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR  
 CC LEUKOCYTES (PMN) GRANULES (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES (BY  
 CC SIMILARITY).

CC -1- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF  
 CC THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE

CC MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE  
 CC RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-

CC AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED  
 CC RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE

CC INGESTED (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: U61270; G1418284;

DR PROSITE: PS00400; LBP\_BPI\_CETP; PARTIAL.

DR PFAM: PF01273; LBP\_BPI\_CETP; 1.

DR HSSP: P17213; 1BPI.

KW ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN.

FT NON\_TER 1

FT SITE 198 203 CLEAVAGE SITES FOR ELASTASE (POTENTIAL).

FT TRANSMEM 327 347 POTENTIAL.

FT CARBOHYD 352 352 POTENTIAL.

SQ SEQUENCE 445 AA; 48837 MW; 5492B04C CRC32;

Query Match

Best Local Similarity 68.0%; Score 34; DB 1; Length 445;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KWKAQKREL 9

|||||  
 DB 80 RWKARKGFI 88

RESULT 14

DPOL\_BPT4  
 ID DPOL\_BPT4 STANDARD; PRT; 898 AA.

AC P04415;

DT 13-AUG-1987 (REL. 05, CREATED)

DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE DNA POLYMERASE (EC 2.7.7.7) (GP43).

GN 43.

OS BACTERIOPHAGE T4.

OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; MYOVIRIDAE;

OC T4-LIKE PHAGES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE: 88227938.

RA SPICER E.K., RUSH J., FUNG C., REHA-KRANTZ L.J., KARAM J.D.,

RA KONIGSBERG W.H.;

RT "Primary structure of T4 DNA polymerase. Evolutionary relatedness to

RT eucaryotic and other procaryotic DNA polymerases.";

RL J. BIOL. CHEM. 263:7478-7485(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE: 89042102.

RA ANDRAKE M., GUILD N., HSU T., GOLD L., TUERK C., KARAM J.;

RT "DNA polymerase of bacteriophage T4 is an autogenous translational

RT repressor.";

RL PROC. NATL. ACAD. SCI. U.S.A. 85:7942-7946(1988).

RN [3]

RP SEQUENCE OF 338-898 FROM N.A.

RX MEDLINE: 88166734.

RA LAMM N., WANG Y., MATHEWS C.K., RUGER W.;

RT "Deoxycytidylate hydroxymethylase gene of bacteriophage T4.

RT Nucleotide sequence determination and over-expression of the gene.";

RL EUR. J. BIOCHEM. 172:553-563(1988).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1-388.

RX MEDLINE: 96292335.

RA WANG J., YU P., LIN T.C., KONIGSBERG W.H., STEITZ T.A.;

RT "Crystal structures of an NH2-terminal fragment of T4 DNA polymerase

RT and its complexes with single-stranded DNA and with divalent metal

RT ions.";

RL BIOCHEMISTRY 35:8110-8119(1996).

CC -1- FUNCTION: THIS POLYMERASE POSSESS TWO ENZYMAIC ACTIVITIES:

CC DNA SYNTHESIS (POLYMERASE) AND AN EXONUCLEOLYTIC ACTIVITY THAT

CC DEGRADES SINGLE STRANDED DNA IN THE 3' TO 5' DIRECTION.

CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -

CC N PYROPHOSPHATE + DNA(N).

CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M10160; G215908;

DR EMBL: X00769; G15382;

DR EMBL: M37159; G215840;

DR PIR: A00717; DJBPT4.

DR PDB: 1NOY; 14-OCT-96.

DR PDB: 1NOZ; 14-OCT-96.

DR PROSITE: PS00116; DNA\_POLYMERASE\_B; 1.

DR PFAM: PF00136; DNA\_pol\_B; 2.

KW TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION; HYDROLASE;  
 KW EXONUCLEASE; DNA-BINDING; 3D-STRUCTURE.

```

FT TRANSMEM 1021 1041 POTENTIAL.
FT TRANSMEM 1091 1111 POTENTIAL.
FT TRANSMEM 1130 1150 POTENTIAL.
FT TRANSMEM 1161 1181 POTENTIAL.
FT TRANSMEM 1223 1243 POTENTIAL.
FT TRANSMEM 1271 1291 POTENTIAL.
FT TRANSMEM 1317 1337 POTENTIAL.
FT CARBOHYD 234 234 POTENTIAL.
FT CARBOHYD 235 235 POTENTIAL.
FT CARBOHYD 865 865 POTENTIAL.
FT CARBOHYD 1430 1430 POTENTIAL.
SQ SEQUENCE 1851 AA; 212168 MW; 601976AI CRC32;

Query Match 70.0%; Score 35; DB 1; Length 1851;
Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAKQRF 8
   ||:||||
Db 424 WRAEKRF 430

RESULT 11
ODO1_HUMAN STANDARD; PRT; 1002 AA.
AC Q02218;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE 2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT PRECURSOR (EC 1.2.4.2)
DE (ALPHA-KETOGLUTARATE DEHYDROGENASE).
GN OGDH.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 92179301.
RA KOIKE K., URATA Y., GOTO S.;
RT "Cloning and nucleotide sequence of the cDNA encoding human 2-oxoglutarate dehydrogenase (lipoamide).";
RL PROC. NATL. ACAD. SCI. U.S.A. 89:1963-1967(1992).
[2]
RP REVISIONS, SEQUENCE FROM N.A.
RX MEDLINE; 95347609.
RA KOIKE K.;
RT "The gene encoding human 2-oxoglutarate dehydrogenase: structural organization and mapping to chromosome 7p13-p14.";
RL GENE 159:261-266(1995).
CC -1- FUNCTION: THE 2-OXOGLUTARATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL CONVERSION OF 2-OXOGLUTARATE TO SUCCINYL-COA & CO(2). IT CONTAINS MULTIPLE COPIES OF 3 ENZYMIC COMPONENTS: 2-OXOGLUTARATE DEHYDROGENASE (E1), DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE (E2) AND LIPOAMIDE DEHYDROGENASE (E3).
CC -1- CATALYTIC ACTIVITY: 2-OXOGLUTARATE + LIPOAMIDE -> S-SUCCINYL-DIHYDROLIPOAMIDE + CO(2).
CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.
CC -1- ENZYME REGULATION: CATABOLITE REPRESSED.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
CC EMBL; D10523; G531241; -
CC EMBL; D32056; G984592; -
CC EMBL; D32057; G984594; -

```

```

DR EMBL; D32058; G984596; -
DR EMBL; D32059; G984598; -
DR EMBL; D32060; G984600; -
DR EMBL; D32060; G984601; -
DR EMBL; D32060; G984602; -
DR EMBL; D32060; G984603; -
DR EMBL; D32061; G984605; -
DR EMBL; D32062; G984607; -
DR EMBL; D32062; G984608; -
DR EMBL; D32062; G984609; -
DR EMBL; D32062; G984610; -
DR EMBL; D32063; G984612; -
DR EMBL; D32063; G984613; -
DR EMBL; D32063; G984614; -
DR EMBL; D32063; G984615; -
DR EMBL; D32063; G984616; -
DR EMBL; D32063; G984617; -
DR EMBL; D32064; G984619; -
DR EMBL; D32064; G984620; -
DR EMBL; D32064; G984621; -
DR PIR; A38234; A38234.
DR MIM; 203740; -
DR PFAM; PF00876; E1_dehydrog; 1.
KW GLYCOLYSIS; OXIDOREDUCTASE; FLAVOPROTEIN; THIAMINE PYROPHOSPHATE; MITOCHONDRION; TRANSIT PEPTIDE.
FT TRANSIT 1 40 MITOCHONDRION.
FT CHAIN 41 1002 ALPHA-KETOGLUTARATE DEHYDROGENASE.
SQ SEQUENCE 1002 AA; 113475 MW; 4851AC95 CRC32;

Query Match 70.0%; Score 35; DB 1; Length 1002;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQRF 8
   ||:||||
Db 271 KWSSEKRF 278

RESULT 12
ATPF_MYCGA STANDARD; PRT; 198 AA.
AC P33256;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE ATP SYNTHASE B CHAIN (EC 3.6.1.34).
GN ATPF.
OS MYCOPLASMA GALLISEPTICUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES; MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92359959.
RA RASMUSSEN O.F., SHIRVAN M.H., MARGALIT H., CHRISTIANSEN C., ROTTEM S.;
RT "Nucleotide sequence, organization and characterization of the atp genes and the encoded subunits of Mycoplasma gallisepticum ATPase.";
RL BIOCHEM. J. 285:881-888(1992).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0) HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SIMILARITY: TO OTHER B SUBUNITS AND ALSO TO B' SUBUNITS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
CC EMBL; X64256; G44298; -

```

```

J. BACTERIOL. 175:4456-4465(1993).
CC -1- FUNCTION: MAY FORM AN ATP-DRIVEN VI POLYSACCHARIDE EXPORT
CC APPARATUS, IN ASSOCIATION WITH THE VEXA, VEXC AND VEXD PROTEINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE ABC-2 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC ENBL: D14156; D1003706;
CC PROSITE; PS00890; ABC2_MEMBRANE; 1.
CC PFAM; PF01061; ABC2_membrane; 1.
CC KW POLYSACCHARIDE TRANSPORT; TRANSPO; TRANSMEMBRANE; INNER MEMBRANE.
CC FT TRANSMEM 38 58 POTENTIAL.
CC FT TRANSMEM 65 85 POTENTIAL.
CC FT TRANSMEM 120 140 POTENTIAL.
CC FT TRANSMEM 149 169 POTENTIAL.
CC FT TRANSMEM 191 211 POTENTIAL.
CC FT TRANSMEM 239 259 POTENTIAL.
CC SEQUENCE 264 AA; 30429 MW; 52811E4C CRC32;
-----
Query Match 72.0%; Score 36; DB 1; Length 264;
Best Local Similarity 66.7%; Pred. No. 5.8;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAKQKRF 9
Db 142 RWEAKQKPF 150
      :|:|||||
-----
RESULT 9
YSX4_CAEL  YSX4_CAEL  STANDARD;  PRT;  434 AA.
ID  YSX4_CAEL  STANDARD;  PRT;  434 AA.
AC  Q10023;
DT  01-OCT-1996 (REL. 34, CREATED)
DT  01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE  01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE  HYPOTHETICAL 50.4 KD PROTEIN T28D9.4 IN CHROMOSOME II.
GN  T28D9.4.
OS  CAENORHABDITIS ELEGANS.
OC  EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC  RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
[1]
RC  SEQUENCE FROM N.A.
RC  STRAIN-BRISTOL N2;
RA  WATERSTON R.;
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC ENBL; U28738; G861265;
CC WORMPEP; T28D9.4; CE02069.
CC HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT  DOMAIN 151 154 POLY-LEU.
FT  TRANSMEM 159 179 POTENTIAL.
FT  TRANSMEM 135 155 POTENTIAL.
FT  TRANSMEM 215 235 POTENTIAL.
FT  TRANSMEM 263 283 POTENTIAL.
-----

```

GN LBP.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C;  
 RX MEDLINE; 97289150.  
 RA LENGACHER S.J., GLAUSER M.P., HEUMANN D.;  
 RA LUEVITCH R.J., LE ROY D., LEE J.D., KRAVCHENKO V.,  
 RT "Reactivity of murine and human recombinant LPS-binding protein (LBP)  
 within LPS and gram negative bacteria.";  
 RL J. INFLAMM. 47:165-172(1995).  
 CC -1- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL  
 CC LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER  
 CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS  
 CC TO INTERACT WITH THE CD14 RECEPTOR.  
 CC -1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X99347; E253895; -  
 CC MGD; MGI:1098776; LBP.  
 DR PROSITE; PS00400; LBP\_BPI\_CETP; 1.  
 DR PFAM; PF01273; LBP\_BPI\_CETP; 1.  
 DR HSSP; P17213; LBP1.  
 KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.  
 FT SIGNAL 1 24  
 FT CHAIN 25 481 LIPOPOLYSACCHARIDE-BINDING PROTEIN.  
 FT CARBOHYD 300 300 POTENTIAL.  
 FT CARBOHYD 355 355 POTENTIAL.  
 SQ SEQUENCE 481 AA; 53312 MW; 973CF58 CRC32;

Query Match 72.08; Score 36; DB 1; Length 481;  
 Best Local Similarity 66.78; Pred. No. 11;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KWAKQKREL 9  
 ||| : |||  
 Db 115 KWVKRKSL 123

RESULT 7  
 PLM1\_PLAFA STANDARD; PRT; 452 AA.  
 AC P39898;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE PLASMEPSIN 1 PRECURSOR (EC 3.4.23.38) (ASPARTIC HEMOGLOBINASE I)  
 DE (PFAPG).  
 OS PLASMODIUM FALCIPARUM.  
 OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDA; PLASMODIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FAF-2;  
 RX MEDLINE; 94147975.  
 RA FRANCIS S.E., GLUZMAN I.Y., OKSMAN A., KNICKERBOCKER A.,  
 RA MUELLER R., BRYANT M.L., SHERMAN D.R., RUSSELL D.G., GOLDBERG D.E.;  
 RT "Molecular characterization and inhibition of a Plasmodium falciparum  
 aspartic hemoglobinase.";  
 RL EMBO J. 13:306-317(1994).  
 RN [2]  
 RP SEQUENCE OF 125-146.  
 RX MEDLINE; 91178457.  
 RA GOLDBERG D.E., SLATER A.F.G., BEAVIS R., CHAIT B., CERAMI A.,

RA HENDERSON G.B.;  
 RT "Hemoglobin degradation in the human malaria pathogen Plasmodium  
 falciparum: a catabolic pathway initiated by a specific aspartic  
 RT protease.";  
 RL J. EXP. MED. 173:961-969(1991).  
 CC -1- FUNCTION: PARTICIPATES IN THE DIGESTION OF THE HOST HEMOGLOBIN.  
 CC INITIAL CLEAVAGE AT THE HINGE REGION OF HEMOGLOBIN, THAN CLEAVES  
 CC AT OTHER SITES, LEADING TO DENATURATION OF THE MOLECULE AND TO  
 CC FURTHER DEGRADATION. OPTIMAL ACTIVITY IS FOUND AT PH 4.5-5.  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 33-PHE-|-LEU-34 BOND IN THE  
 CC ALPHA-CHAIN OF HEMOGLOBIN, LEADING TO DENATURATION OF MOLECULE.  
 CC -1- SUBCELLULAR LOCATION: VACUOLAR. COULD BE FIRST ANCHORED TO THE  
 CC MEMBRANE THROUGH ITS PROPEPTIDE BEFORE BEING RELEASED.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE  
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X75787; G482941; -  
 CC PIR; PT0434; ASP\_PROTEASE; 2.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 DR PFAM; PF00026; asp; 1.  
 DR HSSP; P46925; ISME.  
 KW HYDROLASE; ASPARTYL PROTEASE; GLYCOPROTEIN; ZYMOGEN; SIGNAL.  
 FT SIGNAL 1 124  
 FT PROPEP 125 452 PLASMEPSIN 1.  
 FT CHAIN 157 157 BY SIMILARITY.  
 FT ACT\_SITE 337 337 BY SIMILARITY.  
 FT ACT\_SITE 170 175 BY SIMILARITY.  
 FT DISULFID 372 408 BY SIMILARITY.  
 FT CARBOHYD 184 184 POTENTIAL.  
 FT CARBOHYD 218 218 POTENTIAL.  
 FT CARBOHYD 326 326 POTENTIAL.  
 FT CARBOHYD 440 440 POTENTIAL.  
 SQ SEQUENCE 452 AA; 51461 MW; AD192ADE CRC32;

Query Match 72.08; Score 36; DB 1; Length 452;  
 Best Local Similarity 85.7%; Pred. No. 10;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WKAKRF 8  
 || ||||  
 Db 32 WKIQRF 38  
 RESULT 8  
 VEXB\_SALTI STANDARD; PRT; 264 AA.  
 ID VEXB\_SALTI  
 AC P43109;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE VI POLYSACCHARIDE EXPORT INNER-MEMBRANE PROTEIN VEXB.  
 GN VEXB.  
 OS SALMONELLA TYPHI.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 CC SALMONELLA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GIFU 10007;  
 RX MEDLINE; 93322324.  
 RA HASHIMOTO Y., LI N., YOKOYAMA H., EZAKI T.;  
 RT "Complete nucleotide sequence and molecular characterization of Viab  
 region encoding Vi antigen in Salmonella typhi.";

SQ SEQUENCE 493 AA; 56981 MW; D52DC3DD CRC32;

Query Match 76.0%; Score 38; DB 1; Length 493;  
Best Local Similarity 85.7%; Pred. No. 4.6;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAQKRF 8  
|||||

Db 118 WKAQRRF 124

RESULT 4

WCL\_NEUCR STANDARD; PRT; 1154 AA.  
ID WCL\_NEUCR  
AC Q01371;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE WHITE COLLAR 1 PROTEIN (WCL).  
GN WC-1.  
OS NEUROSPORA CRASSA.  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;  
OC SORDARIALES; SORDARIACEAE; NEUROSPORA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=74-OR23-1A;  
RX MEDLINE; 96203083.  
RA BALLARIO P., VITTORIOSO P., MAGRELLI A., TALORA C., CABIBBO A.,  
RA MACINO G.  
RT "White collar-1, a central regulator of blue light responses in  
RT Neurospora, is a zinc finger protein.";  
RL EMBO J. 15:1650-1657(1996).  
CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR INVOLVED IN LIGHT  
CC REGULATION. BINDS AND AFFECTS BLUE LIGHT REGULATION OF THE AL-3  
CC GENE. WCL AND WC2 INTERACT VIA HOMOLOGOUS PAS DOMAINS, BIND TO  
CC PROMOTERS OF LIGHT REGULATED GENES AND ACTIVATE TRANSCRIPTION.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- SUBUNIT: HETERODIMER OF WCL1 AND WC2 (POTENTIAL).  
CC -1- INDUCTION: BY BLUE LIGHT.  
CC -1- DOMAIN: THE GLUTAMINE-RICH DOMAIN MIGHT FUNCTION IN ACTIVATING  
CC GENE EXPRESSION.  
CC -1- SIMILARITY: TO GATA TRANSCRIPTION FACTORS IN THE ZINC-FINGER  
CC REGION.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; X94300; E256944; -;  
DR PROSITE; PS00344; GATA\_ZN\_FINGER; 1.  
DR PFAM; PF00320; GATA; 1.  
DR PFAM; PF00785; PAC; 1.  
DR HSPSP; P17678; IGAU.  
KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; ZINC-FINGER;  
KW NUCLEAR PROTEIN.  
FT DOMAIN 16 61 GLN-RICH.  
FT ZN\_FING 935 960 GATA-TYPE.  
FT DOMAIN 21 57 POLY-GLN.  
FT DOMAIN 329 333 POLY-PRO.  
SQ SEQUENCE 1154 AA; 125944 MW; F91F27A1 CRC32;

Query Match 74.0%; Score 37; DB 1; Length 1154;  
Best Local Similarity 85.7%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAQKRF 8  
|||||

Db 1138 WKARKRF 1144

RESULT 5

CP18\_DROME STANDARD; PRT; 538 AA.  
ID CP18\_DROME  
AC Q95078; Q27767;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE CYTOCHROME P450 18 (EC 1.14.14.-) (CYPXVIII).  
GN CYP18 OR EIG17-1.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97398464.  
RA BASSETT M.H., MCCARTHY J.L., WATERMAN M.R., SLITER T.J.;  
RT "Sequence and developmental expression of Cyp18, a member of a new  
RT cytochrome P450 family from Drosophila.";  
RL MOL. CELL. ENDOCRINOL. 131:39-49(1997).  
RN [2]  
RP SEQUENCE OF 459-479 FROM N.A.  
RC STRAIN=CANTON-S;  
RX MEDLINE; 94019381.  
RA HURBAN P., THUMMEL C.S.;  
RT "Isolation and characterization of fifteen ecdysone-inducible  
RT Drosophila genes reveal unexpected complexities in ecdysone  
RT regulation.";  
RL MOL. CELL. BIOL. 13:7101-7111(1993).  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).  
CC -1- INDUCTION: BY 20-HYDROXYECDYSONE AT THE TIME OF PUPARIATION.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; U44753; G1517895; -;  
DR EMBL; S66112; G432605; -;  
DR FLYBASE; FBgn010383; Cyp18.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
DR PFAM; PF00067; P450; 1.  
KW OXIDOREDUCTASE; MONOOXYGENASE; TRANSMEMBRANE; HEME.  
FT TRANSMEM 24 44 POTENTIAL.  
FT BINDING 466 466 HEME (BY SIMILARITY).  
FT CONFLICT 478 479 LF->PV (IN REF. 2).  
SQ SEQUENCE 538 AA; 61880 MW; 78A3106C CRC32;

Query Match 72.0%; Score 36; DB 1; Length 538;  
Best Local Similarity 75.0%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAQKRF 9  
|||

Db 140 WKQRRFL 147

RESULT 6

LBP\_MOUSE STANDARD; PRT; 481 AA.  
ID LBP\_MOUSE  
AC O61805;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).



```

CC EMBL; J04739; G179529; ALT_INIT.
DR PIR; A29464; A29464.
DR PIR; A30909; A30909.
DR PIR; A33850; A33850.
DR PDB; 1BPI; 04-SEP-97.
DR MIM; 109195; -.
DR PROSITE; PS00400; LBP_BPI_CETP; 1.
DR PFAM; PF01273; LBP_BPI_CETP; 1.
FT SIGNAL 1 27
FT CHAIN 28 483
KW ANTIBIOTIC; SIGNAL; TRANSMEMBRANE; GLYCOPROTEIN; 3D-STRUCTURE.
FT SIGNAL 1 27
FT CHAIN 28 483
BACTERICIDAL PERMEABILITY-INCREASING
PROTEIN.
FT SITE 236 241
FT CARBOHYD 365 385
FT TRANSMEM 483 AA; 53396 MW; OANCEF65 CRC32;
SQ SEQUENCE 483 AA; 53396 MW; OANCEF65 CRC32;

Query Match 100.0%; Score 50; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKRFL 9
DB 117 KWKAKRFL 125
|||||||

RESULT 2
BPI_BOVIN
ID BPI_BOVIN STANDARD; PRT; 482 AA.
AC P17453;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BACTERICIDAL PERMEABILITY-INCREASING PROTEIN PRECURSOR (BPI).
GN BPI.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RX MEDLINE; 90272418.
RA LEONG S.R., CAMERATO T.;
RT "Nucleotide sequence of the bovine bactericidal permeability
increasing protein (BPI)".
RL NUCLEIC ACIDS RES. 18:3052-3052(1990).
CC -!- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES
OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A
STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE
NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE
GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR
LEUKOCYTES (PMN) GRANULES (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES (BY
SIMILARITY).
CC -!- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF
THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE
MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE
RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-
AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED
RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE
INGESTED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

DR EMBL; X52563; G139; -.
DR PIR; S10180; S10180.
DR PROSITE; PS00400; LBP_BPI_CETP; 1.
DR PFAM; PF01273; LBP_BPI_CETP; 1.
DR HSSP; P17213; LBPI.
KW ANTIBIOTIC; SIGNAL; MEMBRANE; GLYCOPROTEIN.
FT SIGNAL 1 26
FT CHAIN 27 482
BACTERICIDAL PERMEABILITY-INCREASING
PROTEIN.
FT SITE 235 240
FT CARBOHYD 62 62
FT CARBOHYD 303 303
FT CARBOHYD 375 375
FT CARBOHYD 389 389
FT CARBOHYD 463 463
FT SEQUENCE 482 AA; 53432 MW; 26BA9CEFA CRC32;
SQ SEQUENCE 482 AA; 53432 MW; 26BA9CEFA CRC32;

Query Match 78.0%; Score 39; DB 1; Length 482;
Best Local Similarity 66.7%; Pred. No. 2.9;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAKRFL 9
DB 116 KWKAKRFL 124
|||||:|:|

RESULT 3
YDAK_CABEL
ID YDAK_CABEL STANDARD; PRT; 493 AA.
AC P90771; O02645;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE PUTATIVE CYTOCHROME P450 C34B7.3 IN CHROMOSOME I (EC 1.14.-.-).
GN C34B7.3.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA HARRIS B.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA JONES S.J.M.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED
COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; Z83220; E1344512; -.
DR WORMPEP; C34B7.3; CE08567.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR PFAM; PF00067; P450; 1.
KW HYPOTHETICAL PROTEIN; OXIDOREDUCTASE; MONOOXYGENASE; HEME;
KW TRANSMEMBRANE.
FT TRANSMEM 1 21
FT TRANSMEM 60 80
FT TRANSMEM 290 310
FT BINDING 440 440
HEME (BY SIMILARITY).

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:06 ; Search time 71.09 Seconds  
(without alignments)  
3.579 Million cell updates/sec

Title: US-09-124-280A-26

Perfect score: 50

Sequence: 1 KWKAQKRFL 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : Swissprot\_37:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	483	1 BPI_HUMAN	P17213 homo sapien
2	39	78.0	482	1 BPI_BOVIN	P17453 bos taurus
3	38	76.0	493	1 YDAX_CAEEL	P90771 caenorhabdi
4	37	74.0	1154	1 WCL_NEUCR	Q01371 neurospora
5	36	72.0	538	1 CP18_DROME	Q95078 drosophila
6	36	72.0	481	1 LBP_MOUSE	Q61805 mus musculu
7	36	72.0	452	1 PLMI_PLAFA	P39898 plasmodium
8	36	72.0	264	1 VEXE_SALTI	P43109 salmonella
9	36	72.0	434	1 YSXA_CAEEL	Q10023 caenorhabdi
10	35	70.0	1851	1 CCAL_DROME	P91645 drosophila
11	35	70.0	1002	1 ODOI_HUMAN	Q02218 homo sapien
12	34	68.0	198	1 ATPF_MYCGA	P33256 mycoplasma
13	34	68.0	445	1 BPI_RABIT	Q28739 oryctolagus
14	34	68.0	898	1 DPOL_BPT4	P04415 bacterioph
15	34	68.0	285	1 YDH5_SCHPO	Q92350 schizosacch
16	34	68.0	461	1 YUL2_CAEEL	Q19802 caenorhabdi
17	33	66.0	286	1 CBPX_PEA	Q41005 pisum sativ
18	33	66.0	491	1 CPB1_RAT	P00176 rattus norv
19	33	66.0	491	1 CPB2_RAT	P04167 rattus norv
20	33	66.0	500	1 CPJ1_RABIT	P52786 oryctolagus
21	33	66.0	502	1 CPJ2_HUMAN	P51589 homo sapien
22	33	66.0	502	1 CPJ3_RAT	P51590 rattus norv
23	33	66.0	501	1 CPJ6_MOUSE	Q54750 mus musculu
24	33	66.0	2273	1 HFAL_YEAST	P32874 saccharomyc
25	33	66.0	467	1 IRFG_HUMAN	Q14896 homo sapien
26	33	66.0	164	1 RP07_VACCV	P20984 vaccinia vi
27	33	66.0	164	1 RP07_VARV	P33813 variola vir
28	33	66.0	287	1 Y011_MYCGE	P47257 mycoplasma
29	33	66.0	238	1 Y554_AQUAE	Q66829 aquifex aeo
30	32	64.0	489	1 CPJ3_RABIT	P00182 oryctolagus
31	32	64.0	457	1 CPCL_CANFA	P56594 canis fami
32	32	64.0	243	1 GSCB_XENLA	P29454 xenopus lae
33	32	64.0	240	1 GSCB_XENLA	P53546 xenopus lae
34	32	64.0	240	1 GSC_BRARE	P53544 brachydanio
35	32	64.0	245	1 GSC_CHICK	P53545 gallus gall
36	32	64.0	419	1 GSC_DROME	P54366 drosophila
37	32	64.0	256	1 GSC_MOUSE	Q02991 mus musculu
38	32	64.0	403	1 MTIO_SCHPO	Q09878 schizosacch
39	32	64.0	295	1 NADC_YEAST	P43619 saccharomyc
40	32	64.0	389	1 NCAP_CVH22	P15130 human coron
41	32	64.0	350	1 RFBB_RHISN	P55462 rhizobium s
42	32	64.0	496	1 Y018_MYCGE	P47264 mycoplasma
43	32	64.0	954	1 YK15_CAEEL	P49958 caenorhabdi

44 32 64.0 159 1 YMD0\_YEAST Q03713 saccharomyc  
45 31 62.0 66 1 VG5B\_BPPH2 P03683 bacterioph

#### ALIGNMENTS

RESULT 1

BPI\_HUMAN  
ID BPI\_HUMAN STANDARD; PRT; 483 AA.  
AC P17213;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE BACTERICIDAL PERMEABILITY-INCREASING PROTEIN PRECURSOR (BPI) (CAP 57).  
GN BPI.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-64.  
RX MEDLINE: 89255455.  
RA GRAY P.W., FLAGGS G., LEONG S.R., GUMINA R.J., WEISS J., OOI C.E.,  
RA ELSBACH P.;  
RT Cloning of the cDNA of a human neutrophil bactericidal protein.  
RT Structural and functional correlations.;  
RL J. BIOL. CHEM. 264:9505-9509(1989).  
RN [2]  
RP SEQUENCE OF 28-42.  
RX MEDLINE: 88033057.  
RA OOI C.E., WEISS J., ELSBACH P., FRANGIONE B., MANNION B.;  
RT "A 25-kDa NH2-terminal fragment carries all the antibacterial  
RT activities of the human neutrophil 60-kDa  
RT bactericidal/permeability-increasing protein.";  
RL J. BIOL. CHEM. 262:14891-14894(1987).  
RN [3]  
RP SEQUENCE OF 28-47.  
RX MEDLINE: 89315847.  
RA GABAY J.E., SCOTT R.W., CAMPANELLI D., GRIFFITH J., WILDE C.,  
RA MARRA M.N., SEGER M., NATHAN C.F.;  
RT "Antibiotic proteins of human polymorphonuclear leukocytes.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:5610-5614(1989).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
RX MEDLINE: 97334442.  
RA BEAMER L.J., CARROLL S.F., EISENBERG D.;  
RT "Crystal structure of human BPI and two bound phospholipids at 2.4-A  
RT resolution.";  
RL SCIENCE 276:1861-1864(1997).  
CC -!- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES  
CC OF GRAM-NEGATIVE BACTERIA: THIS SPECIFICITY MAY BE EXPLAINED BY A  
CC STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE  
CC NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE  
CC GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR  
CC LEUKOCYTES (PMN) GRANULES.  
CC -!- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES.  
CC -!- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF  
CC THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE  
CC MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE  
CC RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-  
CC AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED  
CC RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE  
CC INGESTED.  
CC -!- SIMILARITY: BELONGS TO THE BPI/CTP/GBP/PLTP FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).



```

F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match          66.0%; Score 33; DB 1; Length 491;
Best Local Similarity 62.5%; Pred. No. 66;
Matches 5; Conservative 2; Mismatches 1; Indels 0;

Qy      1 KWKAKRRF 8
      : : : : :
Db      120 RWKALRRF 127

Search completed: September 7, 1999, 23:22:41
Job time: 948 sec

```

A:Residues: 338-898 <LAM>  
A:Cross-references: GB:M37159; GB:M28192; NID:g215839; PID:g215840  
C:Comment: This enzyme is required for both initiation and maintenance of viral DNA replication  
C:Genetics: 43  
A:Gene: 43  
A:Map position: 27.201-29.895  
C:Superfamily: Phage T4 DNA-directed DNA polymerase  
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 68.0%; Score 34; DB 1; Length 898;  
Best Local Similarity 71.4%; Pred. No. 79;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAQKRF 8  
|||:|:  
DB 699 WKAKRY 705

RESULT 11  
S24335  
H+-transporting ATP synthase (EC 3.6.1.34) chain b - Mycoplasma gallisepticum (SGC3)  
C:Species: Mycoplasma gallisepticum  
C:Date: 23-Apr-1993 #sequence\_revision 23-Apr-1993 #text\_change 04-Sep-1998  
C:Accession: S24335  
R:Rasmussen, O.F.; Shirvan, M.H.; Margalit, H.; Christiansen, C.; Rottem, S.  
Biochem. J. 285, 881-888, 1992  
A:Title: Nucleotide sequence, organization and characterization of the atp genes and the  
A:Reference number: S24331  
A:Accession: S24335  
A:Molecule type: DNA  
A:Residues: 1-198 <RAS>  
A:Cross-references: EMBL:X64256; NID:g44294; PID:g44298  
C:Genetics:  
A:Gene: atpF  
A:Genetic code: SGC3  
C:Superfamily: H+-transporting ATP synthase chain B  
C:Keywords: hydrolase

Query Match 68.0%; Score 34; DB 2; Length 198;  
Best Local Similarity 75.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 WKAQKRF 9  
|||:|:  
DB 67 WKPKRF 74

RESULT 12  
D70140  
hypothetical protein BB0325 - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
C:Accession: D70140  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
son, D.; Peterson, J.; Kervavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: D70140  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-369 <KLE>  
A:Cross-references: GB:AE001139; GB:AE000783; NID:g2688225; PID:g2688231; TIGR:BB0325  
A:Experimental source: strain B31

Query Match 68.0%; Score 34; DB 2; Length 369;  
Best Local Similarity 50.0%; Pred. No. 32;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAQKRFL 9  
|||:|:  
DB 324 WRAERRFI 331

RESULT 13  
S76420  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
C:Accession: S76420  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76420  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-156 <KAN>  
A:Cross-references: EMBL:D90915; GB:AB001339; NID:gl653604; PID:d1019282; PID:gl65363  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 68.0%; Score 34; DB 2; Length 156;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWAQKRFL 9  
|||:|:  
DB 52 RWEAQRGL 60

RESULT 14  
F69326  
hypothetical protein AF0614 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
C:Accession: F69326  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kalne, B.P.; Sykes,  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
A:Reference number: A69250; MUID:98049343  
A:Accession: F69326  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-93 <KLE>  
A:Cross-references: GB:AE001062; GB:AE000782; NID:g2689385; PID:g2650010; TIGR:AF0614

Query Match 68.0%; Score 34; DB 2; Length 93;  
Best Local Similarity 55.6%; Pred. No. 7.7;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWAQKRFL 9  
|||:|:  
DB 75 EWKSKRFI 83

RESULT 15  
O4RTPB  
cytochrome P450 2B1 - rat  
N:Alternate names: cytochrome P450 b; cytochrome P450, phenobarbital-inducible  
N:Contains: unspecified monooxygenase (EC 1.14.14.1)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 18-Aug-1982 #sequence\_revision 17-May-1996 #text\_change 05-Mar-1999

A:Cross-references: GB:D14156; NID:g426443; PID:d1003706; PID:g426450

C:Genetics:

A:Gene: vexB

Query Match 72.0%; Score 36; DB 2; Length 264;

Best Local Similarity 66.7%; Pred. No. 9.3;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRF 9

||:|||||

Db 142 RWEAKPFL 150

RESULT 7

S41717

aspartic hemoglobinase (EC 3.4.23.-) - Plasmodium falciparum

C:Species: Plasmodium falciparum

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Jan-1999

C:Accession: S41717; PT0434

R:Francis, S.E.; Gluzman, I.Y.; Oksman, A.; Knickerbocker, A.; Mueller, R.; Bryant, M.L.

EMBO J. 13, 306-317, 1994

A:Title: Molecular characterization and inhibition of a Plasmodium falciparum aspartic h

A:Reference number: S41717; MUID:94147975

A:Accession: S41717

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-452 <FRA>

A:Cross-references: GB:X75787; NID:g482940; PID:g482941

R:Goldberg, D.E.; Slater, A.F.G.; Beavis, R.; Chait, B.; Cerami, A.; Henderson, G.B.

J. Exp. Med. 173, 961-969, 1991

A:Title: Hemoglobin degradation in the human malaria pathogen Plasmodium falciparum: a c

A:Reference number: PT0434; MUID:91178457

A:Accession: PT0434

A:Molecule type: protein

A:Residues: 125-146 <GOL>

A>Note: This sequence reveals 9 of 22 residues identical to the most specific mammalian

C:Comment: This enzyme plays a key role in the hemoglobinolytic pathway.

C:Keywords: aspartic proteinase; hydrolase

Query Match 72.0%; Score 36; DB 2; Length 452;

Best Local Similarity 85.7%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKQAQKRF 8

||:|||||

Db 32 WKIQKRF 38

RESULT 8

A38234

oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2) precursor - human

N:Alternate names: 2-oxoglutarate:lipoamide 2-oxidoreductase; alpha-ketoglutarate dehyd

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 13-Nov-1998

C:Accession: A38234

R:Koike, K.; Urata, Y.; Goto, S.

Proc. Natl. Acad. Sci. U.S.A. 89, 1963-1967, 1992

A:Title: Cloning and nucleotide sequence of the cDNA encoding human 2-oxoglutarate dehyd

A:Reference number: A38234; MUID:92179301

A:Accession: A38234

A:Molecule type: mRNA

A:Residues: 1-1003 <KOI>

A:Cross-references: GB:D10523

A:Experimental source: fetal liver

A>Note: sequence extracted from NCBI backbone (NCBIP:87352)

C:Superfamily: oxoglutarate dehydrogenase (lipoamide); thiamine pyrophosphate-binding de

C:Keywords: mitochondrion; oxidoreductase; thiamin pyrophosphate; tricarboxylic acid cyc

F:1-40/Domain: transit peptide (mitochondrion) #status predicted <TNP>

F:41-1003/Product: oxoglutarate dehydrogenase (lipoamide) #status predicted <MAT>

F:408-453/Domain: thiamine pyrophosphate-binding domain homology <TPB>

Query Match 70.0%; Score 35; DB 2; Length 1003;

Best Local Similarity 62.5%; Pred. No. 57;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRF 8

||:|||||

Db 271 KWSSEKRF 278

RESULT 9

E70458

translation releasing factor RF-2 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Dec-1998

C:Accession: E70458

R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

V. Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666

A:Accession: E70458

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-373 <AQF>

A:Cross-references: GB:AE000758; NID:g2984111; PID:g2984119; GB:AE000657

A:Experimental source: strain VF5

C:Genetics:

A:Gene: prfB

C:Superfamily: translation releasing factor

Query Match 70.0%; Score 35; DB 2; Length 373;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQK 6

||||||

Db 361 KWKAQK 366

RESULT 10

DJBPT4

DNA-directed DNA polymerase (EC 2.7.7.7) - phage T4

N:Alternate names: gp 43

C:Species: phage T4

A>Note: host Escherichia coli

C:Date: 17-Mar-1987 #sequence\_revision 30-Jun-1987 #text\_change 22-Jan-1999

C:Accession: JS0791; PU0006; A28165; JU0097; A00717

R:Spicer, E.K.; Konigsberg, W.H.

in Bacteriophage T4, Mathews, C.K., Kutter, E.M., Mosig, G., and Berget, P.B., eds.,

A:Title: Organization and structure of four T4 genes coding for DNA replication prote

A:Reference number: A04303

A:Accession: JS0791

A:Molecule type: DNA

A:Residues: 1-898 <SPI>

A:Accession: PU0006

A:Molecule type: protein

A:Residues: 1-22 <SP2>

R:Spicer, E.K.; Rush, J.; Fung, C.; Reha-Krantz, L.J.; Karam, J.D.; Konigsberg, W.H.

J. Biol. Chem. 263, 7478-7486, 1988

A:Title: Primary structure of T4 DNA polymerase.

A:Reference number: A28165; MUID:88227938

A:Accession: A28165

A:Molecule type: DNA

A:Residues: 1-898 <SP3>

A:Cross-references: GB:M10160; GB:J02510; GB:X00769; NID:g2947028; PID:g215908

R:Lamm, N.; Wang, Y.; Mathews, C.K.; Rueger, W.

Eur. J. Biochem. 172, 553-563, 1988

A:Title: Deoxycytidylate hydroxymethylase gene of bacteriophage T4: nucleotide sequen

A:Reference number: JF0071; MUID:88166734

A:Accession: JU0097

A:Molecule type: DNA

Query Match 100.0%; Score 50; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.036;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
|||||  
Db 121 KWKAQKRFL 129

RESULT 2  
S43383  
bactericidal/permeability-increasing protein - synthetic  
C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in *Escherichia coli*  
C:Date: 20-Oct-1994 #sequence\_revision 15-Feb-1996 #text\_change 15-Feb-1996  
C:Accession: S43383  
R:Q1, S.Y.; Li, Y.; O'Connor, C.D.  
Biochem. J. 238, 711-718, 1994  
A:Title: The region around residue 115 of human bactericidal/permeability-increasing protein  
of a gene coding for the active domain and characterization of recombinant proteins.  
A:Reference number: S43383  
A:Accession: S43383  
A:Molecule type: DNA  
A:Residues: 1-250 <QIS>

Query Match 100.0%; Score 50; DB 4; Length 250;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
|||||  
Db 91 KWKAQKRFL 99

RESULT 3  
S10180  
bactericidal permeability-increasing protein precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 08-Sep-1997  
C:Accession: S10180  
R:Leon, S.R.; Camerato, T.  
Nucleic Acids Res. 18, 3052, 1990  
A:Title: Nucleotide sequence of the bovine bactericidal permeability increasing protein  
A:Reference number: S10180; MUID:90272418  
A:Accession: S10180  
A:Molecule type: mRNA  
A:Residues: 1-482 <LEO>  
A:Cross-references: EMBL:X52563; NID:gl138; PID:gl139  
C:Superfamily: lipopolysaccharide-binding protein  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-482/Product: bactericidal permeability increasing protein #status predicted <MAT>

Query Match 78.0%; Score 39; DB 2; Length 482;  
Best Local Similarity 66.7%; Pred. No. 4.6;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
|||||  
Db 116 KWKAQKRFL 124

RESULT 4  
S69206  
regulator protein white collar 1 - *Neurospora crassa*  
C:Species: *Neurospora crassa*  
C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 10-Oct-1997  
C:Accession: S69206  
R:Ballario, P.; Vittorioso, P.; Magrelli, A.; Talora, C.; Cabibbo, A.; Macino, G.  
EMBO J. 15, 1650-1657, 1996  
A:Title: White collar-1, a central regulator of blue light responses in *Neurospora*, is a  
A:Reference number: S69206; MUID:96203083

Query Match 100.0%; Score 50; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.036;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
|||||  
Db 121 KWKAQKRFL 129

RESULT 2  
S43383  
bactericidal/permeability-increasing protein - synthetic  
C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in *Escherichia coli*  
C:Date: 20-Oct-1994 #sequence\_revision 15-Feb-1996 #text\_change 15-Feb-1996  
C:Accession: S43383  
R:Q1, S.Y.; Li, Y.; O'Connor, C.D.  
Biochem. J. 238, 711-718, 1994  
A:Title: The region around residue 115 of human bactericidal/permeability-increasing protein  
of a gene coding for the active domain and characterization of recombinant proteins.  
A:Reference number: S43383  
A:Accession: S43383  
A:Molecule type: DNA  
A:Residues: 1-250 <QIS>

Query Match 100.0%; Score 50; DB 4; Length 250;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
|||||  
Db 91 KWKAQKRFL 99

RESULT 3  
S10180  
bactericidal permeability-increasing protein precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 08-Sep-1997  
C:Accession: S10180  
R:Leon, S.R.; Camerato, T.  
Nucleic Acids Res. 18, 3052, 1990  
A:Title: Nucleotide sequence of the bovine bactericidal permeability increasing protein  
A:Reference number: S10180; MUID:90272418  
A:Accession: S10180  
A:Molecule type: mRNA  
A:Residues: 1-482 <LEO>  
A:Cross-references: EMBL:X52563; NID:gl138; PID:gl139  
C:Superfamily: lipopolysaccharide-binding protein  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-482/Product: bactericidal permeability increasing protein #status predicted <MAT>

Query Match 78.0%; Score 39; DB 2; Length 482;  
Best Local Similarity 66.7%; Pred. No. 4.6;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
|||||  
Db 116 KWKAQKRFL 124

RESULT 4  
S69206  
regulator protein white collar 1 - *Neurospora crassa*  
C:Species: *Neurospora crassa*  
C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 10-Oct-1997  
C:Accession: S69206  
R:Ballario, P.; Vittorioso, P.; Magrelli, A.; Talora, C.; Cabibbo, A.; Macino, G.  
EMBO J. 15, 1650-1657, 1996  
A:Title: White collar-1, a central regulator of blue light responses in *Neurospora*, is a  
A:Reference number: S69206; MUID:96203083

A:Accession: S69206  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1154 <BAL>  
A:Cross-references: EMBL:X94300; NID:gl279576; PID:e256944; PID:gl480115  
C:Genetics:  
A:Introns: 967/3  
C:Superfamily: GATA-type zinc finger homology  
F:932-991/Domain: GATA-type zinc finger homology <GZF>

Query Match 74.0%; Score 37; DB 2; Length 1154;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAKRKF 8  
|||||  
Db 1138 WKAKRKF 1144

RESULT 5  
S67642  
probable arsenical pump-driving ATPase (EC 3.6.1.-) YDL100c - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein D2371  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 18-Sep-1998  
C:Accession: S67642; S67419  
R:Ballesta, J.P.G.; Renacha, M.; Soler-Mira, A.; Jimenez, A.; Garcia-Cantalejo, J.M.;  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67629  
A:Accession: S67642  
A:Molecule type: DNA  
A:Residues: 1-354 <BAL>  
A:Cross-references: EMBL:Z74148; NID:gl431137; PID:e253031; PID:gl431138; MIPS:YDL100  
A:Experimental source: strain S288C  
R:Boskovic, J.; Saiz, J.E.; Soler-Mira, A.; Garcia-Cantalejo, J.; Revuelta, J.L.; Jim  
submitted to the EMBL Data Library, February 1996  
A:Reference number: S67406  
A:Accession: S67419  
A:Molecule type: DNA  
A:Residues: 1-354 <BOS>  
A:Cross-references: EMBL:X95644; NID:gl199535; PID:e223235; PID:gl199549  
C:Genetics:  
A:Map position: 4L  
A:Note: YDL100c  
C:Superfamily: arsenical pump-driving ATPase  
C:Keywords: ATP; hydrolase; P-loop  
F:25-32/Region: nucleotide-binding motif A (P-loop)

Query Match 72.0%; Score 36; DB 2; Length 354;  
Best Local Similarity 55.6%; Pred. No. 13;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
|||||  
Db 291 RWKMKKYL 299

RESULT 6  
G36892  
VI polysaccharide capsule transporter VexB - *Salmonella typhi*  
C:Species: *Salmonella typhi*  
C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 09-Sep-1997  
C:Accession: G36892  
R:Hashimoto, Y.; Li, N.; Yokoyama, H.; Ezaki, T.  
J. Bacteriol. 175, 4456-4465, 1993  
A:Title: Complete nucleotide sequence and molecular characterization of *viab* region e  
A:Reference number: A36892; MUID:93322324  
A:Accession: G36892  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <HAS>

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:40 ; Search time 105.14 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-26

Perfect score: 50

Sequence: 1 KWKAQKREL 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database: PIR\_60.\*

1: .pir1.\*  
2: .pir2.\*  
3: .pir3.\*  
4: .pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	487	2	A30909
2	50	100.0	250	4	S43383
3	39	78.0	482	2	S10180
4	37	74.0	1154	2	S69206
5	36	72.0	354	2	S87642
6	36	72.0	264	2	G36892
7	36	72.0	452	2	S41717
8	35	70.0	1003	2	A38234
9	35	70.0	373	2	E70458
10	34	68.0	898	1	DJBPN4
11	34	68.0	198	2	S24335
12	34	68.0	369	2	D70140
13	34	68.0	156	2	S76420
14	34	68.0	93	2	F69326
15	33	66.0	491	1	O4RTPB
16	33	66.0	491	1	O4RTP2
17	33	66.0	164	1	RNV219
18	33	66.0	501	2	A40938
19	33	66.0	367	2	S19172
20	33	66.0	184	2	F36848
21	33	66.0	286	2	S62370
22	33	66.0	2123	2	S55089
23	33	66.0	459	2	JC6520
24	33	66.0	238	2	A70350
25	33	66.0	287	2	B64201
26	33	66.0	462	2	T01732
27	33	66.0	1704	2	S71363
28	33	66.0	337	2	C70406
29	32	64.0	489	1	O4RBP3
30	32	64.0	389	1	VH1H2E
31	32	64.0	457	2	A37222
32	32	64.0	252	2	A54677
33	32	64.0	256	2	A42768
34	32	64.0	245	2	A47539
35	32	64.0	243	2	B42768
36	32	64.0	243	2	I51424
37	32	64.0	389	2	S08031
38	32	64.0	295	2	S56302
39	32	64.0	687	2	B64126

## ALIGNMENTS

RESULT 1

A30909

bactericidal/permeability-increasing protein precursor - human  
N:Alternate names: 55K bactericidal protein  
C:Species: Homo sapiens (man)  
C>Date: 18-Apr-1989 #sequence revision 18-Apr-1989 #text change 17-Mar-1999  
C:Accession: A33850; B54136; A29464; A43600; A49716; A30909  
R:Gray, P.W.; Flagg, G.; Leong, S.R.; Gumina, R.J.; Weiss, J.; Ooi, C.E.; Elsbach, P.  
J. Biol. Chem. 264, 9505-9509, 1989  
A:Title: Cloning of the cDNA of a human neutrophil bactericidal protein. Structural a  
A:Reference number: A33850; MUID:89255455  
A:Accession: A33850  
A:Molecule type: mRNA  
A:Residues: 1-487 <GRA>  
R:Wild, C.G.; Seilhamer, J.J.; McGrogan, M.; Ashton, N.; Snable, J.L.; Lane, J.C.; L  
J. Biol. Chem. 269, 17411-17416, 1994  
A:Title: Bactericidal/permeability-increasing protein and lipopolysaccharide (LPS)-bi  
A:Reference number: A54136; MUID:94292492  
A:Accession: B54136  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-374, 'L', 376-487 <WIL>  
A:Experimental source: HL-60 cells  
A:Note: sequence extracted from NCBI backbone (NCBIP:149855)  
R:Ooi, C.E.; Weiss, J.; Elsbach, P.; Frangione, B.; Mannion, B.  
J. Biol. Chem. 262, 14891-14894, 1987  
A:Title: A 25-kDa amino-terminal fragment carries all the antibacterial activities of  
A:Reference number: A29464; MUID:88033057  
A:Accession: A29464  
A:Molecule type: protein  
A:Residues: 32-51 <OOI>  
A:Experimental source: neutrophils  
R:Wasiluk, K.R.; Skubitz, K.M.; Gray, B.H.  
Infect. Immun. 59, 4193-4200, 1991  
A:Title: Comparison of granule proteins from human polymorphonuclear leukocytes which  
A:Reference number: A43600; MUID:92040097  
A:Accession: A43600  
A:Molecule type: protein  
A:Residues: 32-52, 'R', <WAS>  
R:Little, R.G.; Kelner, D.N.; Lim, E.; Burke, D.J.; Conlon, P.J.  
J. Biol. Chem. 269, 1865-1872, 1994  
A:Title: Functional domains of recombinant bactericidal/permeability increasing prote  
A:Reference number: A49716; MUID:94124531  
A:Accession: A49716  
A:Molecule type: protein  
A:Residues: 32-130;132-141;143-165;202-215,'E',217-225 <LIT>  
C:Comment: The bactericidal/permeability-increasing protein (BPI) is a 60 kD membrane  
which is specific for gram-negative bacteria. BPI has a high affinity for lipopolysac  
between BPI and an LPS-binding protein from liver and cholesterol ester transfer prot  
C:Genetics:  
A:Gene: GDB:BPI  
A:Cross-references: GDB:131572; OMIM:109195  
A:Map position: 20q11.23-20q12  
C:Superfamily: lipopolysaccharide-binding protein  
C:Keywords: antibacterial; cytotoxin; glycoprotein; heparin binding; neutrophil  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-487/Product: bactericidal permeability-increasing protein #status predicted <MAT  
F:32-51/Region: bactericidal #status predicted  
F:380/Binding site: carbohydrate (Asn) (covalent) #status predicted

40	32	64.0	149	2	S74398	hypothetical prote
41	32	64.0	1414	2	B70674	probable mbtB prot
42	32	64.0	427	2	I64201	helicase (moti) ho
43	32	64.0	443	2	T01182	protein kinase hom
44	32	64.0	403	2	S62531	hypothetical prote
45	32	64.0	159	2	S49749	probable membrane



**THIS PAGE BLANK (USPTO)**

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,611A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/273,401  
FILING DATE: 11-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/125,651  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 32251  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "BPI.17"  
US-08-311-611A-18

Query Match 100.0%; Score 50; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFL 9  
|||||  
DB 6 KWKAQRFL 14

RESULT 15  
US-08-311-611A-19  
Sequence 19, Application US/08311611A  
Patent No. 5523288  
GENERAL INFORMATION:  
APPLICANT: Cohen, Jonathan  
APPLICANT: Kung, Ada H.C.  
APPLICANT: Lambert, Jr., Lewis H.  
TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
TITLE OF INVENTION: Infection by Administration of  
TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 227  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,611A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/273,401  
FILING DATE: 11-JUL-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/125,651  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 32251  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "BPI.18"  
US-08-311-611A-19

Query Match 100.0%; Score 50; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFL 9  
|||||  
DB 6 KWKAQRFL 14

Search completed: September 7, 1999, 23:07:46  
Job time: 1744 sec

```
; Sequence 16, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311.611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "BPI.15"
; US-08-311-611A-16

Query Match 100.0%; Score 50; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWAQKREL 9
Db 6 KWAQKREL 14

RESULT 13
US-08-311-611A-17
; Sequence 17, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311.611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "BPI.16"
; US-08-311-611A-17

Query Match 100.0%; Score 50; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWAQKREL 9
Db 6 KWAQKREL 14

RESULT 14
US-08-311-611A-18
; Sequence 18, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "BPI.58"  
US-08-311-611A-9

Query Match 100.0%; Score 50; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0057;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
Db 7 KWKAQKRFL 15

RESULT 10  
US-08-311-611A-10  
; Sequence 10, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "BPI.65 oxidized"  
US-08-311-611A-10

Query Match 100.0%; Score 50; DB 1; Length 17;

Best Local Similarity 100.0%; Pred. No. 0.0061;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
Db 7 KWKAQKRFL 15

RESULT 11  
US-08-311-611A-11  
; Sequence 11, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "BPI.3"  
US-08-311-611A-11

Query Match 100.0%; Score 50; DB 1; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.0092;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
Db 18 KWKAQKRFL 26

RESULT 12  
US-08-311-611A-16

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/311.611A  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/273,401  
;; FILING DATE: 11-JUL-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/125,651  
;; FILING DATE: 22-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sharp, Jeffrey S.  
;; REGISTRATION NUMBER: 31,879  
;; REFERENCE/DOCKET NUMBER: 32251  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/474-6300  
;; TELEFAX: 312/474-0448  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 15 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "BPI.2"  
;; US-08-311-611A-7

Query Match 100.0%; Score 50; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRFL 9  
Db 6 KWKAQKRFL 14

RESULT 8  
US-08-311-611A-8  
;; Sequence 8, Application US/08311611A  
;; Patent No. 5523288  
;; GENERAL INFORMATION:  
;; APPLICANT: Cohen, Jonathan  
;; APPLICANT: Kung, Ada H.C.  
;; APPLICANT: Lambert, Jr., Lewis H.  
;; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
;; TITLE OF INVENTION: Infection by Administration of  
;; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
;; TITLE OF INVENTION:  
;; NUMBER OF SEQUENCES: 227  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
;; STREET: 6300 Sears Tower, 233 South Wacker Drive  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60606-6402  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/311.611A  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/273,401  
;; FILING DATE: 11-JUL-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/125,651  
;; FILING DATE: 22-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:

;; NAME: Sharp, Jeffrey S.  
;; REGISTRATION NUMBER: 31,879  
;; REFERENCE/DOCKET NUMBER: 32251  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/474-6300  
;; TELEFAX: 312/474-0448  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "BPI.8"  
;; US-08-311-611A-8

Query Match 100.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0037;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRFL 9  
Db 1 KWKAQKRFL 9

RESULT 9  
US-08-311-611A-9  
;; Sequence 9, Application US/08311611A  
;; Patent No. 5523288  
;; GENERAL INFORMATION:  
;; APPLICANT: Cohen, Jonathan  
;; APPLICANT: Kung, Ada H.C.  
;; APPLICANT: Lambert, Jr., Lewis H.  
;; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
;; TITLE OF INVENTION: Infection by Administration of  
;; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
;; TITLE OF INVENTION:  
;; NUMBER OF SEQUENCES: 227  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
;; STREET: 6300 Sears Tower, 233 South Wacker Drive  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60606-6402  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/311.611A  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/273,401  
;; FILING DATE: 11-JUL-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/125,651  
;; FILING DATE: 22-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sharp, Jeffrey S.  
;; REGISTRATION NUMBER: 31,879  
;; REFERENCE/DOCKET NUMBER: 32251  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/474-6300  
;; TELEFAX: 312/474-0448  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 16 amino acids

QY 1 KWKAQRFL 9  
|||||  
Db 121 KWKAQRFL 129

## RESULT 5

US-08-414-924-2  
; Sequence 2, Application US/08414924  
; Patent No. 5494896  
; GENERAL INFORMATION:  
; APPLICANT: Hansbrough, John F.  
; TITLE OF INVENTION: Method of Treating Conditions  
; TITLE OF INVENTION: Associated With Burn Injuries  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/414,924  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 27129/32297  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-414-924-2

Query Match 100.0%; Score 50; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFL 9  
|||||  
Db 121 KWKAQRFL 129

## RESULT 6

US-08-311-611A-6  
; Sequence 6, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "Domain II"  
US-08-311-611A-6

Query Match 100.0%; Score 50; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFL 9  
|||||  
Db 26 KWKAQRFL 34

## RESULT 7

US-08-311-611A-7  
; Sequence 7, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

APPLICANT: Grinna, Lynn S  
TITLE OF INVENTION: Stable Bactericidal/Permeability-  
TITLE OF INVENTION: Increasing Protein Products and Pharmaceutical  
TITLE OF INVENTION: Compositions Containing the Same  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/013,801  
FILING DATE: 02 FEB 1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Thomas C.  
REGISTRATION NUMBER: P-36,989  
REFERENCE/DOCKET NUMBER: 27129/30911  
TELEPHONE: 312/346-5750  
TELEFAX: 312/346-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-013-801-2

Query Match 100.0%; Score 50; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.13; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFL 9  
DB 121 KWKAQRFL 129  
RESULT 3  
US-08-072-063-2  
Sequence 2, Application US/08072063  
Patent No. 5439807  
GENERAL INFORMATION:  
APPLICANT: Theofan, Georgia  
APPLICANT: Grinna, Lynn S  
APPLICANT: Horwitz, Arnold  
TITLE OF INVENTION: Bp1-Immunoglobulin Fusion Proteins  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/072,063  
FILING DATE: 19930519

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers Thomas C.  
REGISTRATION NUMBER: 36,989  
REFERENCE/DOCKET NUMBER: 30659  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-072-063-2  
Query Match 100.0%; Score 50; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KWKAQRFL 9  
DB 121 KWKAQRFL 129  
RESULT 4  
US-08-212-132-2  
Sequence 2, Application US/08212132  
Patent No. 5447913  
GENERAL INFORMATION:  
APPLICANT: Little, Roger G.  
APPLICANT: Ammons, William Steve  
TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-  
TITLE OF INVENTION: Increasing Protein Dimer Products  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/212,132  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 27129/31735  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-212-132-2

Query Match 100.0%; Score 50; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:46 ; Search time 82.37 Seconds  
(without alignments)  
1.078 Million cell updates/sec

Title: US-09-124-280A-26  
Perfect score: 50  
Sequence: 1 KWKAQKREL 9

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	487	1	US-08-030-644-2
2	50	100.0	487	1	US-08-013-801-2
3	50	100.0	487	1	US-08-072-063-2
4	50	100.0	487	1	US-08-212-132-2
5	50	100.0	487	1	US-08-414-924-2
6	50	100.0	35	1	US-08-311-611A-6
7	50	100.0	15	1	US-08-311-611A-7
8	50	100.0	10	1	US-08-311-611A-8
9	50	100.0	16	1	US-08-311-611A-9
10	50	100.0	17	1	US-08-311-611A-10
11	50	100.0	27	1	US-08-311-611A-11
12	50	100.0	15	1	US-08-311-611A-16
13	50	100.0	15	1	US-08-311-611A-17
14	50	100.0	15	1	US-08-311-611A-18
15	50	100.0	15	1	US-08-311-611A-19
16	50	100.0	15	1	US-08-311-611A-20
17	50	100.0	15	1	US-08-311-611A-29
18	50	100.0	15	1	US-08-311-611A-32
19	50	100.0	15	1	US-08-311-611A-49
20	50	100.0	15	1	US-08-311-611A-51
21	50	100.0	24	1	US-08-311-611A-52
22	50	100.0	29	1	US-08-311-611A-53
23	50	100.0	20	1	US-08-311-611A-54
24	50	100.0	25	1	US-08-311-611A-55
25	50	100.0	20	1	US-08-311-611A-57
26	50	100.0	20	1	US-08-311-611A-58
27	50	100.0	26	1	US-08-311-611A-65
28	50	100.0	17	1	US-08-311-611A-66
29	50	100.0	17	1	US-08-311-611A-68
30	50	100.0	487	1	US-08-311-611A-69
31	50	100.0	24	1	US-08-311-611A-70
32	50	100.0	487	1	US-08-311-611A-146
33	50	100.0	30	1	US-08-311-611A-149
34	50	100.0	20	1	US-08-311-611A-150
35	50	100.0	20	1	US-08-311-611A-151
36	50	100.0	14	1	US-08-311-611A-164
37	50	100.0	12	1	US-08-311-611A-227
38	50	100.0	487	1	US-08-173-968-2
39	50	100.0	221	1	US-08-173-968-3

40 50 100.0 487 1 US-08-232-527-2  
41 50 100.0 35 1 US-08-372-783-6  
42 50 100.0 15 1 US-08-372-783-7  
43 50 100.0 10 1 US-08-372-783-8  
44 50 100.0 16 1 US-08-372-783-9  
45 50 100.0 487 3 PCT-US95-09262-206

ALIGNMENTS

RESULT 1  
US-08-030-644-2  
; Sequence 2, Application US/08030644  
; Patent No. 5348942  
; GENERAL INFORMATION:  
; APPLICANT: Little, Roger G. II  
; APPLICANT: Gazzano-Santoro, Helene  
; APPLICANT: Parent, James Brian  
; TITLE OF INVENTION: Therapeutic uses of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing Protein Products  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/030.644  
; FILING DATE: 19930312  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 31229  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-030-644-2

Query Match 100.0%; Score 50; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.13; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKREL 9  
|||||

DB 121 KWKAQKREL 129

RESULT 2  
US-08-013-801-2  
; Sequence 2, Application US/08013801  
; Patent No. 5420019  
; GENERAL INFORMATION:  
; APPLICANT: Theofan, Georgia  
; APPLICANT: Horwitz, Arnold  
; APPLICANT: Burke, David  
; APPLICANT: Baltayan, Manik



**THIS PAGE BLANK (USPTO)**

Search completed: September 7, 1999, 22:49:57  
Job time: 7878 sec

CC rheumatoid arthritis, and for treating gram-negative bacterial infection.  
 CC The peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipopolysaccharide in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipopolysaccharide or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 SQ Sequence 20 AA;

Query Match 100.0%; Score 50; DB 1; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.01;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRF 9

Db 1 KWKAQKRF 9

RESULT 14

R63727 ID R63727 standard; peptide; 15 AA.

AC R63727;

DT 03-MAY-1995 (first entry)

DE BPI derived peptide, BPI.9.

KW Human; bactericidal/permeability-increasing protein; BPI; heparin;

KW binding agent; neutralisation; anti-coagulant effect; inhibition;

KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;

KW contraception; malignant; tumour cell; inflammatory disease; T-cell;

KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;

KW lipopolysaccharide; circulation; compromised immune response; microbe;

KW macrophage; activation; lymphokine; decontaminating; Helicobacter;

KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;

KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.

OS Homo sapiens.

PN WO9420532-A.

PD 15-SEP-1994.

PF 11-MAR-1994; U02465.

PR 12-MAR-1993; US-030644.

PR 15-JUL-1993; US-093202.

PR 14-JAN-1994; US-183222.

PA (XOMA ) XOMA CORP.

PI Little RG.

DR WPI; 94-302964/37.

PT New human bactericidal permeability increasing peptides - derived

PT from the functional domains of BPI and having BPI activities such

PT as bactericidal activity

PS Claim 8; Page 158; 254pp; English.

CC The sequences given in R63682-750, R62087-100 and R62491-500 are

CC peptides derived from human bactericidal/permeability-increasing

CC protein (BPI). The sequences given in R63727-35 are derived from

CC positions 65-99 of BPI and may form dimers or trimers. Peptides

CC such as these may be used as heparin binding agents, for neutralising

CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.

CC associated with ocular retinopathy, for inhibiting endothelial cell

CC proliferation, for contraception, for inhibiting malignant tumour cell

CC proliferation, for treating a chronic inflammatory disease state, eg.

CC rheumatoid arthritis, and for treating gram-negative bacterial infection.

CC The peptides may also be used for treating a subject suffering from the

CC adverse effects of the presence of lipopolysaccharide in the circulation,

CC eg. a compromised immune response to microbes or tumour cells due to

CC inhibition of macrophage activation by T-cell lymphokines or increased

CC production of a cytokine, for decontaminating a fluid containing

CC lipopolysaccharide or for treating a disease associated with Helicobacter

CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.

CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B

CC or cefamandole nafate or LBP protein products. The peptides are pref.

CC prepared by solid phase synthesis.

CC Sequence 15 AA;

Query Match 100.0%; Score 50; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.008;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRF 9

Db 6 KWKAQKRF 14

RESULT 15

R63730

ID R63730 standard; peptide; 20 AA.

AC R63730;

DT 03-MAY-1995 (first entry)

DE BPI derived peptide, BPI.46.

KW Human; bactericidal/permeability-increasing protein; BPI; heparin;

KW binding agent; neutralisation; anti-coagulant effect; inhibition;

KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;

KW contraception; malignant; tumour cell; inflammatory disease; T-cell;

KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;

KW lipopolysaccharide; circulation; compromised immune response; microbe;

KW macrophage; activation; lymphokine; decontaminating; Helicobacter;

KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;

KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.

OS Homo sapiens.

PN WO9420532-A.

PD 15-SEP-1994.

PF 11-MAR-1994; U02465.

PR 12-MAR-1993; US-030644.

PR 15-JUL-1993; US-093202.

PR 14-JAN-1994; US-183222.

PA (XOMA ) XOMA CORP.

PI Little RG.

DR WPI; 94-302964/37.

PT New human bactericidal permeability increasing peptides - derived

PT from the functional domains of BPI and having BPI activities such

PT as bactericidal activity

PS Claim 8; Page 159; 254pp; English.

CC The sequences given in R63682-750, R62087-100 and R62491-500 are

CC peptides derived from human bactericidal/permeability-increasing

CC protein (BPI). The sequences given in R63727-35 are derived from

CC positions 65-99 of BPI and may form dimers or trimers. Peptides

CC such as these may be used as heparin binding agents, for neutralising

CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.

CC associated with ocular retinopathy, for inhibiting endothelial cell

CC proliferation, for contraception, for inhibiting malignant tumour cell

CC proliferation, for treating a chronic inflammatory disease state, eg.

CC rheumatoid arthritis, and for treating gram-negative bacterial infection.

CC The peptides may also be used for treating a subject suffering from the

CC adverse effects of the presence of lipopolysaccharide in the circulation,

CC eg. a compromised immune response to microbes or tumour cells due to

CC inhibition of macrophage activation by T-cell lymphokines or increased

CC production of a cytokine, for decontaminating a fluid containing

CC lipopolysaccharide or for treating a disease associated with Helicobacter

CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.

CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B

CC or cefamandole nafate or LBP protein products. The peptides are pref.

CC prepared by solid phase synthesis.

CC Sequence 20 AA;

Query Match 100.0%; Score 50; DB 1; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.01;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRF 9

Db 11 KWKAQKRF 19

```

RESULT 11
R57901
ID R57901 standard; Protein; 487 AA.
AC R57901;
DE 29-MAR-1995 (first entry)
KW Bactericide BPI.
KW BPI; bactericidal-permeability-increasing protein; bactericide;
KW fusion protein; Gram-negative bacterium; infection.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1..31
FT /label= Sig_peptide
FT
FT WO9418323-A.
FT 18-AUG-1994.
PD 18-AUG-1994.
PF 02-FEB-1993; U01235.
PR 02-FEB-1993; US-013801.
PA (XOMA ) XOMA CORP
PI Baitalan M, Burke D, Grinna L, Horwitz A, Theofan G;
DR WPI: 94-279744/34.
DR N-PSDB: Q67270.
PT Bactericidal-permeability-increasing protein analogs with
PT Cys132 or Cys135 replaced - also fusion proteins and C-terminally
PT truncated forms, have increased stability and are useful to treat
PT gram-negative bacterial infection
PS Disclosure: Page 52-54: 77pp; English.
CC Analogs, fusion proteins and C-terminal truncated forms of human BPI
CC are claimed that show improved resistance to dimerization and
CC aduct formation.
SQ Sequence 487 AA;

Query Match 100.0%; Score 50; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFL 9
Db 121 KWKAQRFL 129
|||||

RESULT 12
R62370
ID R62370 standard; Protein; 487 AA.
AC R62370;
DE 27-APR-1995 (first entry)
DE Recombinant bactericidal/permeability-increasing protein (rBPI-50).
KW N-terminal; recombinant; bactericidal/permeability protein; rBPI;
KW signal protein; human; rBPI-23; binding assay; heparin-binding; domain;
KW Limulus amoebocyte lysate; LAL; inhibition; bactericidal activity;
KW chronic inflammatory disease; arthritis; psoriasis; Crohn's disease;
KW inflammatory bowel disease; asthma.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1..31
FT /note= "Signal peptide"
FT protein 32..487
FT /note= "mature protein"
FT domain 17..55
FT /note= "LAL inhibiting domain"
FT binding_site 21..55
FT /note= "Heparin binding site"
FT binding_site 65..107
FT /note= "Heparin binding site"
FT domain 73..99
FT /note= "LAL inhibiting domain"
FT active_site 85..99
FT /note= "Bactericidally active fragment"
FT domain 137..163
FT /note= "LAL inhibiting domain"
FT binding_site 137..171
FT /note= "Heparin binding site"
FT
FT WO9420128-A.
FT 15-SEP-1994.

```

```

PF 11-MAR-1994; U02401.
PR 12-MAR-1993; US-030644.
PR 15-JUL-1993; US-093202.
PA (XOMA ) XOMA CORP.
PI Gazzano-Santoro H, Little RG, Parent JB;
DR WPI: 94-302679/37.
DR N-PSDB: Q73382.
PT Uses of bactericidal permeability increasing protein product - to
PT neutralise anticoagulant effect of heparin, to inhibit
PT angiogenesis or cell proliferation, as contraceptive and to reduce
PT inflammation
PS Disclosure: Page 43-45; 78pp; English.
CC This sequence represents the N-terminal of recombinant bactericidal/
CC permeability (rBPI) protein. This expression product encoding the 31
CC residue signal protein and the first 199 amino acids of the N-terminal
CC of mature human BPI has a molecular weight of 23 kD and is designated
CC rBPI-23. In binding assays rBPI-23 was shown to have three separate
CC functional domains with heparin-binding activity, at least three major
CC domains with significant Limulus amoebocyte lysate (LAL) inhibition and
CC one region with bactericidal activity. BPI, or active fragments of it
CC may be used to treat and prevent the effects of chronic inflammatory
CC disease states such as arthritis, psoriasis, inflammatory bowel disease,
CC Crohn's disease and asthma.
SQ Sequence 487 AA;

Query Match 100.0%; Score 50; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFL 9
Db 121 KWKAQRFL 129
|||||

RESULT 13
R63731
ID R63731 standard; peptide; 20 AA.
AC R63731;
DE 03-MAY-1995 (first entry)
DE BPI derived peptide, BPI.47.
KW Human; bactericidal/permeability-increasing protein; BPI; heparin;
KW binding agent; neutralisation; anti-coagulant effect; inhibition;
KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;
KW contraception; malignant; tumour cell; inflammatory disease; T-cell;
KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;
KW lipaarabinomannan; circulation; compromised immune response; microbe;
KW macrophage; activation; lymphokine; decontaminating; Helicobacter;
KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;
KW gentamicin; polymyxin B; cefamandole nateate; LBP protein.
OS Homo sapiens.
PN WO9420532-A.
PD 15-SEP-1994.
PF 11-MAR-1994; U02465.
PR 12-MAR-1993; US-030644.
PR 15-JUL-1993; US-093202.
PR 14-JAN-1994; US-183222.
PA (XOMA ) XOMA CORP.
PI Little RG;
DR WPI: 94-302964/37.
PT New human bactericidal permeability increasing peptides - derived
PT from the functional domains of BPI and having BPI activities such
PT as bactericidal activity
PS Claim 8; page 159; 254pp; English.
CC The sequences given in R63682-750, R62087-100 and R62491-500 are
CC peptides derived from human bactericidal/permeability-increasing
CC protein (BPI). The sequences given in R63727-35 are derived from
CC positions 65-99 of BPI and may form dimers or trimers. Peptides
CC such as these may be used as heparin binding agents, for neutralising
CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.
CC associated with ocular retinopathy, for inhibiting endothelial cell
CC proliferation, for contraception, for inhibiting malignant tumour cell
CC proliferation, for treating a chronic inflammatory disease state, eg.

```

FT /note= "Mature rBPI(191)-IgG fusion protein"  
 PN WO9323434-A.  
 PD 25-NOV-1993.  
 PF 19-MAY-1993; U04754.  
 PR 19-MAY-1992; US-885911.  
 PA (XOMA ) XOMA CORP.  
 PI Grinna LS, Horwitz A, Theofan G;  
 DR WPI; 93-386485/48.  
 DR N-PSDB: Q52488.  
 PT New fusion proteins for treating bacterial infections - comprising  
 PT a bactericidal-permeability-increasing protein and an immunoglobulin  
 PT heavy chain constant domain  
 PS Claim 20: Page 48-50; 75pp; English.  
 CC This sequence is encoded by the plasmid pING4512. pING4512 was  
 CC constructed using the primer sequence given in Q52480. Plasmid  
 CC pING4503 was cut with AlwNI, the ends made blunt with T4 DNA  
 CC polymerase, and the DNA then cut with SalI. The resultant approx.  
 CC 700bp SalI/blunt fragment contains the 30 bp 5' untranslated region  
 CC and the DNA encoding the signal peptide and first 191 amino acids of  
 CC bactericidal/permeability-increasing protein (BPI). IgG heavy chain  
 CC (HC) sequences, which also included the hinge region of the IgG HC  
 CC were PCR amplified from plasmid pMB27 using these primers. Plasmids  
 CC such as pING4512 may be used to produce fusion proteins which are  
 CC useful for the treatment of gram negative bacterial infections and  
 CC their sequelae including endotoxin related shock and conditions  
 CC associated with it, such as disseminated intravascular coagulation,  
 CC anemia, thrombocytopenia, leukopenia, adult respiratory distress  
 CC syndrome, renal failure, hypotension, fever and metabolic acidosis.  
 CC Proving BPI or a fragment of it, as part of the fusion with an  
 CC immunoglobulin heavy chain constant region provides the potential  
 CC advantages of Fc receptor binding, bivalent binding to lipopoly-  
 CC saccharide complement binding and increasing placental transfer.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 50; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
 Db 121 KWKAQKRFL 129

RESULT 9  
 ID R43886 standard; Protein; 487 AA.  
 AC R43886;  
 DE 01-JUN-1994 (first entry)  
 DE Human bactericidal/permeability-increasing protein.  
 KW Bactericidal/permeability-increasing protein;  
 KW endotoxin binding protein.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT protein 32..487  
 FT /label= human BPI protein  
 FT misc\_difference 1..31  
 FT /label= signal peptide  
 PN WO9323540-A.  
 PD 25-NOV-1993.  
 PF 19-MAY-1993; U04752.  
 PR 19-MAY-1992; US-885501.  
 PA (XOMA ) XOMA CORP.  
 PI Grinna LS;  
 DR WPI; 93-386572/48.  
 DR P-PSDB: Q52268.  
 PT Recovering endotoxin binding protein from transformed cell  
 PT cultures - by retention on cation exchange material and elution,  
 PT giving simpler recovery and higher yields, for treating bacterial  
 PT infections  
 PS Disclosure; Page 27-28; 38pp; English.  
 CC Human bactericidal/permeability increasing protein (an endotoxin  
 CC binding protein) is useful for treatment of bacterial infections.

SQ Sequence 487 AA;

Query Match 100.0%; Score 50; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
 Db 121 KWKAQKRFL 129

RESULT 10

ID R63687 standard; peptide; 35 AA.  
 AC R63687;  
 DE BPI derived peptide, domain II.  
 DT 03-MAY-1995 (first entry)  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipopolysaccharide; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 OS WO9420532-A.  
 PD 15-SEP-1994.  
 PF 11-MAR-1993; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 13-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.  
 PA (XOMA ) XOMA CORP.  
 PI Little RG;  
 DR WPI; 94-302964/37.  
 PT New human bactericidal permeability increasing peptides - derived  
 PT from the functional domains of BPI and having BPI activities such  
 PT as bactericidal activity  
 PS Claim 6; Page 157; 254pp; English.  
 CC The sequences given in R63682-750, R62087-100 and R62491-500 are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC protein (BPI). The sequences given in R63687-726 are derived from  
 CC positions 65-99 of BPI. Peptides such as these may be used as  
 CC heparin binding agents, for neutralising the anti-coagulant effect  
 CC of heparin, for inhibiting angiogenesis, eg. associated with ocular  
 CC retinopathy, for inhibiting endothelial cell proliferation, for  
 CC contraception, for inhibiting malignant tumour cell proliferation,  
 CC for treating a chronic inflammatory disease state, eg. rheumatoid  
 CC arthritis, and for treating gram-negative bacterial infection. The  
 CC peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipopolysaccharide in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipopolysaccharide or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 SQ Sequence 35 AA;

Query Match 100.0%; Score 50; DB 1; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 0.018;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
 Db 26 KWKAQKRFL 34

CC The hsp1 holoprotein was isolated from human promyelocytic leukemia  
 CC cells (HL-60) and BPI cDNA was prep'd. The N-terminal fragments of  
 CC BPI exhibit all the anti-bacterial and membrane permeability-  
 CC increasing functions contained in the complete molecule (as well as  
 CC the LPS-neutralising function) but are substantially smaller. In  
 CC addition, the N-terminal fragment, as well as the C-terminal  
 CC fragment(s) are capable of inhibiting the prodn. and/or release of  
 CC tumour necrosis factor (TNF) triggered by LPS in whole blood. The  
 CC inventors claim the DNA sequence comprising bps 721-1491 of Q25593.  
 CC They also claim the following fragments of MATURE hBPI: 1-99 (I);  
 CC 1-203 (II); 200-456 (III); 204-456 (IV). NB: The above fragments  
 CC are labeled with respect to the first AA of the mature sequence,  
 CC which corresp. to residue 32 of R24777.  
 CC Sequence 487 AA;  
 SQ

Query Match 100.0%; Score 50; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
 |||||  
 DB 121 KWKAQKRFL 129

## RESULT 6

ID R33776 standard; protein; 368 AA.  
 AC R33776;  
 DE 22-JUL-1993 (first entry)  
 DE BPI from p337.  
 KW Bacterial/permeability increasing protein; lipopolysaccharide; LPS;  
 KW inhibitor; endotoxin-related shock; disseminated intravascular  
 KW coagulation; anaemia; thrombocytopenia; ARDS; renal failure;  
 KW endotoxemia; adult respiratory distress syndrome.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..32  
 FT region /note= "signal peptide"  
 FT 32..368  
 FT /note= "mature protein"  
 PN W09305797-A.  
 PD 01-APR-1993.  
 PF 28-SEP-1992; U08234.  
 PR (INCY-) INCYTE PHARM INC.  
 PA Marra MN, Scott RW, Snable JL, Wilde CG;  
 PI WPI; 93-117241/14.  
 DR Compns. contg. a bactericidal and permeability increasing  
 PT protein - solubilised in a lipid carrier to maintain stably an  
 PT active form  
 PS Disclosure; Fig 6; 53pp; English.  
 CC The human bactericidal/permeability increasing protein may be used  
 CC in a compn. solubilised in a lipid carrier. The BPI binds  
 CC lipopolysaccharides and inhibits the immunostimulatory and toxic  
 CC activities of LPS in vitro and in vivo. The lipid compsn. stably  
 CC maintains the BPI in a sol. active form. The compsn. can be used  
 CC for detecting or quantifying LPS in a sample, coating a surgical  
 CC tool, coating an implantable invasive device, decontaminating a  
 CC fluid contg. LPC and treating endotoxin related -shock,  
 CC -disseminated intravascular coagulation, -anaemia, -thrombocytopenia,  
 CC -adult respiratory stress syndrome or -renal failure, or for  
 CC preventing endotoxaemia.  
 CC See also R33777, R33729-31.  
 SQ Sequence 368 AA;

Query Match 100.0%; Score 50; DB 1; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
 |||||

DB 121 KWKAQKRFL 129

## RESULT 7

ID R33777 standard; protein; 243 AA.  
 AC R33777;  
 DE 22-JUL-1993 (first entry)  
 DE BPI from p212.  
 KW Bacterial/permeability increasing protein; lipopolysaccharide; LPS;  
 KW inhibitor; endotoxin-related shock; disseminated intravascular  
 KW coagulation; anaemia; thrombocytopenia; ARDS; renal failure;  
 KW endotoxemia; adult respiratory distress syndrome.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..32  
 FT region /note= "signal peptide"  
 FT 32..243  
 FT /note= "mature protein"  
 PN W09305797-A.  
 PD 01-APR-1993.  
 PF 28-SEP-1992; U08234.  
 PR 27-SEP-1991; US-766566.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW, Snable JL, Wilde CG;  
 DR WPI; 93-117241/14.  
 DR Compns. contg. a bactericidal and permeability increasing  
 PT protein - solubilised in a lipid carrier to maintain stably an  
 PT active form  
 PS Disclosure; Fig 7; 53pp; English.  
 CC The human bactericidal/permeability increasing protein may be used  
 CC in a compn. solubilised in a lipid carrier. The BPI binds  
 CC lipopolysaccharides and inhibits the immunostimulatory and toxic  
 CC activities of LPS in vitro and in vivo. The lipid compsn. stably  
 CC maintains the BPI in a sol. active form. The compsn. can be used  
 CC for detecting or quantifying LPS in a sample, coating a surgical  
 CC tool, coating an implantable invasive device, decontaminating a  
 CC fluid contg. LPC and treating endotoxin related -shock,  
 CC -disseminated intravascular coagulation, -anaemia, -thrombocytopenia,  
 CC -adult respiratory stress syndrome or -renal failure, or for  
 CC preventing endotoxaemia.  
 CC See also R33776, R33729-31.  
 SQ Sequence 243 AA;

Query Match 100.0%; Score 50; DB 1; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFL 9  
 |||||  
 DB 121 KWKAQKRFL 129

## RESULT 8

ID R44420 standard; protein; 487 AA.  
 AC R44420;  
 DE 03-JUN-1994 (first entry)  
 DE PING4512 encoded fusion protein.  
 KW Polymase chain reaction; primer; amplify; PCR; plasmid; PING4512;  
 KW PING4503; bactericidal/permeability-increasing protein; BPI; IgG;  
 KW heavy; chain; PMB27; fusion protein; gram negative; renal failure;  
 KW bacterial infection; endotoxin related shock; metabolic acidosis;  
 KW disseminated intravascular coagulation; anaemia; thrombocytopenia;  
 KW leukopenia; adult respiratory distress syndrome; hypotension; fever;  
 KW constant region; lipopolysaccharide complement binding;  
 KW placental transfer; FC receptor binding.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT peptide 1..31  
 FT region /note= "signal peptide"  
 FT 32..487  
 FT protein

```

FT region 1. .31
PN WO9203535-A. /label= residues -31 to -1
PD 05-MAR-1992.
PF 13-AUG-1991; U05758.
PR 13-AUG-1990; US-567016.
PA (INCY-) INCYTE PHARM INC.
PI Marra MN, Scott RW;
DR WPI: 92-096879/12.
PT New recombinant BPI protein - specifically binds to and inhibits
PT endotoxin lethality, for prevention, treatment and diagnosis of
PT endo-toxaemia and endo-toxic shock
PS Disclosure: Fig 13; 108pp; English.
CC The BPI binds endotoxin in the presence of serum and plasma and,
CC unlike other known endotoxin binding proteins such as LBP
CC (lipopolysaccharide binding protein), BPI inhibits the
CC immunostimulatory and toxic activities of endotoxin both in vitro
CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic
CC and prophylactic treatment of endotoxin-related disorders including
CC endotoxaemia and endotoxic shock.
SQ Sequence 368 AA;

Query Match 100.0%; Score 50; DB 1; Length 368;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAKRFL 9
Db 121 KWKAKRFL 129

RESULT 3
R21843
ID R21843 standard; Protein; 396 AA.
AC R21843;
DE 22-JUL-1992 (first entry)
DE Sequence of CHO-BPI a biologically active variant of BPI, a human
DE 57kd protein which binds to the outer membrane of susceptible gram
DE negative bacteria (BPI).
KW Endotoxin binding protein; diagnosis; therapy; endotoxaemia;
KW endotoxic shock.
OS Homo sapiens.
PN WO9203535-A.
PD 05-MAR-1992.
PF 13-AUG-1991; U05758.
PR 13-AUG-1990; US-567016.
PA (INCY-) INCYTE PHARM INC.
PI Marra MN, Scott RW;
DR WPI: 92-096879/12.
PT New recombinant BPI protein - specifically binds to and inhibits
PT endotoxin lethality, for prevention, treatment and diagnosis of
PT endo-toxaemia and endo-toxic shock
PS Claim 5; Fig 24; 108pp; English.
CC The BPI binds endotoxin in the presence of serum and plasma and,
CC unlike other known endotoxin binding proteins such as LBP
CC (lipopolysaccharide binding protein), BPI inhibits the
CC immunostimulatory and toxic activities of endotoxin both in vitro
CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic
CC and prophylactic treatment of endotoxin-related disorders including
CC endotoxaemia and endotoxic shock.
SQ Sequence 396 AA;

Query Match 100.0%; Score 50; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAKRFL 9
Db 90 KWKAKRFL 98

```

```

RESULT 4
R21844
ID R21844 standard; Protein; 456 AA.
AC R21844;
DE 22-JUL-1992 (first entry)
DE Sequence of BPI (DP linkage), a biologically active variant of BPI, a
DE 57kd protein which binds to the outer membrane of susceptible gram
DE negative bacteria (BPI).
KW Endotoxin binding protein; diagnosis; therapy; endotoxaemia;
KW endotoxic shock.
OS Homo sapiens.
PN WO9203535-A.
PD 05-MAR-1992.
PF 13-AUG-1991; U05758.
PR 13-AUG-1990; US-567016.
PA (INCY-) INCYTE PHARM INC.
PI Marra MN, Scott RW;
DR WPI: 92-096879/12.
PT New recombinant BPI protein - specifically binds to and inhibits
PT endotoxin lethality, for prevention, treatment and diagnosis of
PT endo-toxaemia and endo-toxic shock
PS Claim 52; Fig 25; 108pp; English.
CC The BPI binds endotoxin in the presence of serum and plasma and,
CC unlike other known endotoxin binding proteins such as LBP
CC (lipopolysaccharide binding protein), BPI inhibits the
CC immunostimulatory and toxic activities of endotoxin both in vitro
CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic
CC and prophylactic treatment of endotoxin-related disorders including
CC endotoxaemia and endotoxic shock.
SQ Sequence 456 AA;

Query Match 100.0%; Score 50; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAKRFL 9
Db 90 KWKAKRFL 98

RESULT 5
R24777
ID R24777 standard; Protein; 487 AA.
AC R24777;
DE 02-JAN-1992 (first entry)
DE Sequence of human bactericidal/permeability-increasing protein
DE (hBPI) derived from human promyelocytic leukemia cells (HL-60).
KW Bactericidal; endotoxin inhibitor; anti-microbial.
OS Homo sapiens.
PN WO9209621-A.
DR WPI: 92-217016/26.
PT Antibacterial fragments of bactericidal-permeability-increasing
PT protein - for treating Gram negative bacterial infections esp. in
PT immuno suppressed patients, with endotoxin activity
PS Claim 16; Fig 5; 63pp; English.

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:49:57 ; Search time 135.78 Seconds  
(without alignments)  
1.570 Million cell updates/sec

Title: US-09-124-280A-26

Perfect score: 50

Sequence: 1 KWAKQKRL 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	50	100.0	487	1	P94636	Human bactericidal
2	50	100.0	368	1	R21840	Sequence of p337,
3	50	100.0	396	1	R21843	Sequence of CHO-BP
4	50	100.0	456	1	R21844	Sequence of BPI (D
5	50	100.0	487	1	R24777	Sequence of human
6	50	100.0	368	1	R33776	BPIP from p337. Co
7	50	100.0	243	1	R33777	BPIP from p212. Co
8	50	100.0	487	1	R44420	PING4512 encoded f
9	50	100.0	487	1	R43886	Human bactericidal
10	50	100.0	35	1	R63687	BPI derived peptid
11	50	100.0	487	1	R57901	Bactericide BPI. B
12	50	100.0	487	1	R62370	Recombinant bacter
13	50	100.0	20	1	R63731	BPI derived peptid
14	50	100.0	15	1	R63727	BPI derived peptid
15	50	100.0	20	1	R63730	BPI derived peptid
16	50	100.0	24	1	R63729	BPI derived peptid
17	50	100.0	21	1	R63728	BPI derived peptid
18	50	100.0	26	1	R63735	BPI derived peptid
19	50	100.0	24	1	R62495	BPI derived peptid
20	50	100.0	29	1	R62496	BPI derived peptid
21	50	100.0	24	1	R62497	BPI derived peptid
22	50	100.0	15	1	R63706	BPI derived peptid
23	50	100.0	15	1	R63709	BPI derived peptid
24	50	100.0	15	1	R63715	BPI derived peptid
25	50	100.0	17	1	R63720	BPI derived peptid
26	50	100.0	10	1	R63689	BPI derived peptid
27	50	100.0	16	1	R63690	BPI derived peptid
28	50	100.0	17	1	R63891	BPI derived peptid
29	50	100.0	27	1	R63692	BPI derived peptid
30	50	100.0	15	1	R63693	BPI derived peptid
31	50	100.0	15	1	R63694	BPI derived peptid
32	50	100.0	15	1	R63695	BPI derived peptid
33	50	100.0	15	1	R63696	BPI derived peptid
34	50	100.0	15	1	R63697	BPI derived peptid
35	50	100.0	487	1	R62344	Recombinant 23 kD
36	50	100.0	15	1	R63688	BPI derived peptid
37	50	100.0	487	1	R67998	Holoprotein rBPI50
38	50	100.0	487	1	R71290	Human bactericidal
39	50	100.0	368	1	R71291	Human bactericidal
40	50	100.0	243	1	R71292	Human bactericidal
41	50	100.0	9	1	R71797	Peptide neutralisi
42	50	100.0	10	1	R71798	Peptide neutralisi
43	50	100.0	199	1	R68925	BPI(1-159)/lipopol

44 50 100.0 197 1 R68929 Lipopolysaccharide  
45 50 100.0 11 1 W10353 Peptide C90-99 for

#### ALIGNMENTS

##### RESULT 1

P94636

ID P94636 standard; protein; 487 AA.

AC P94636;

DT 29-JUN-1990 (first entry)

DE Human bactericidal/permeability-increasing (BPI) holoprotein

KW Polypeptide fragment; bacteraemia; sepsis; gram-negative bacteria;

KW bactericidal/permeability-increasing (BPI) holoprotein;

KW mammalian myeloid blood cells; polymorphonuclear leukocytes (PMN);

KW human promyelocytic leukaemia cells (HL-60).

OS Homo Sapiens.

FH Key

FT peptide Location/Qualifiers

FT protein 1..31

FT region 32..487

FT region 153..155

FT region /note="potential glycosylation site"

FT region 379..381

FT region /note="potential glycosylation site"

PN W08901486-A.

PD 23-FEB-1989.

PF 09-AUG-1988; U02700.

PR 05-AUG-1988; US-228035, US-084335.

PA (U)NY New York Univ.

PI Elsbach P, Weiss J;

DR WPI: 89-068849/09.

DR N-PSDB: N92766

PT Bacterial-permeability increasing holoprotein fragments -

PT used for treatment of diseases caused by gram negative bacteria

PT such as bacteraemia or sepsis.

PS Disclosure: ; 41pp; English.

PT The purified, isolated cDNA which encodes this sequence was isolated

CC from human promyelocytic leukaemia cells (HL-60) and is also claimed.

CC Biologically active BPI protein fragments can be used for the

CC treatment of mammals suffering from diseases caused by gram-negative

CC bacteria such as bacteraemia or sepsis. They have good selectivity and

CC lack of cytotoxicity towards cells other than gram-negative bacteria.

CC The fragments may be used in a prep. further comprising rifampicin,

CC penicillin or cephalosporin. The BPI holoproteins used in prodn. of

CC the fragments may be obtd. from mammalian myeloid blood cells, eg

CC polymorphonuclear leukocytes (PMN), or by recombinant DNA techniques.

CC BPI fragments are prepd. by incubating the holoprotein under protein

CC cleaving conditions or using recombinant DNA techniques.

SQ Sequence 487 AA;

Query Match 100.0%; Score 50; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWAKQKRL 9

DB 121 KWAKQKRL 129

##### RESULT 2

R21840

ID R21840 standard; protein; 368 AA.

AC R21840;

DT 22-JUL-1992 (first entry)

DE Sequence of p337, a biologically active fragment of the human

DE 57Kd protein which binds to the outer membrane of susceptible gram

DE negative bacteria (BPI).

KW Endotoxin binding protein; diagnosis; therapy; endotoxemia;

KW endotoxin shock.

OS Homo sapiens.

FH Key

FT peptide Location/Qualifiers



**THIS PAGE BLANK (USPTO)**

Best Local Similarity 75.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
|||||:  
DB 133 KELTKVKI 140

## RESULT 14

O73967 PRELIMINARY; PRT; 67 AA.  
AC O73967;  
DT 01-NOV-1998 (TReMBLrel. 08, Created)  
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE 67AA LONG HYPOTHETICAL PROTEIN.  
GN PH5014.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-OT3;  
RX MEDLINE; 98344137.  
RA KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSIYAMA A., NAGAI Y.,  
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
RA KIKUCHI H.;  
RT "Complete sequence and gene organization of the genome of a  
hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
DR EMBL; AP000002; BAA29491.1;  
SQ SEQUENCE 67 AA; 7995 MW; 030B838C CRC32;

Query Match 81.1%; Score 30; DB 1; Length 67;  
Best Local Similarity 62.5%; Pred. No. 77;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
|||||:  
DB 11 KELERLKV 18

## RESULT 15

O58391 PRELIMINARY; PRT; 499 AA.  
AC O58391;  
DT 01-AUG-1998 (TReMBLrel. 07, Created)  
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
DE 499AA LONG HYPOTHETICAL PHENYLALANYL-TRNA SYNTHETASE SUBUNIT ALPHA  
DE CHAIN.  
GN PH0658.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-OT3;  
RX MEDLINE; 98344137.  
RA KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSIYAMA A., NAGAI Y.,  
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
RA KIKUCHI H.;  
RT "Complete sequence and gene organization of the genome of a  
hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
DR EMBL; AP000003; BAA29749.1;  
DR PFAM; PF01409; tRNA-synt\_2d; 1.  
KW Aminoacyl-tRNA synthetase.

SQ SEQUENCE 499 AA; 57532 MW; ABD19506 CRC32;

Query Match 81.1%; Score 30; DB 1; Length 499;  
Best Local Similarity 87.5%; Pred. No. 4.3e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
|||||:  
DB 164 KELRRKI 171

Search completed: September 7, 1999, 22:47:34  
Job time: 7954 sec

RN SEQUENCE FROM N.A.  
 RP MATTHEWS L.;  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KESHAU J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z78545; CAB01765.1; -;  
 SQ SEQUENCE 150 AA; 17428 MW; E6AD3D44 CRC32;

Query Match 83.8%; Score 31; DB 5; Length 150;  
 Best Local Similarity 75.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELRIKI 8  
 Db 55 KELQRLKI 62  
 |||||:|  
 |||||:|

RESULT 11  
 O96246 PRELIMINARY; PRT; 980 AA.  
 AC O96246;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DE 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE HYPOTHETICAL 118.9 KD PROTEIN.  
 GN PFB0765W.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
 RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
 RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
 RT falciparum";  
 RL Science 282:1126-1132(1998).  
 DR EMBL; AE001417; AAC71942.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 980 AA; 118857 MW; EA2372E8 CRC32;

Query Match 83.8%; Score 31; DB 5; Length 980;  
 Best Local Similarity 85.7%; Pred. No. 5e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELRIKI 7  
 Db 570 KELRLK 576  
 |||||:|  
 |||||:|

RESULT 12  
 Q33862 PRELIMINARY; PRT; 645 AA.  
 ID Q33862  
 AC Q33862;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE FUMARATE REDUCTASE FLAVOPROTEIN SUBUNIT.  
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Ascaridida;  
 OC Ascaridoidea; Ascarididae; Ascaris.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95258075.  
 RA KURAMUCHI T., HIRAWAKE H., KOJIMA S., TAKAMIYA S., FURUSHIMA R.,  
 RA AOKI T., KOMUNTECKI R., KITA K.;  
 RT "Sequence comparison between the flavoprotein subunit of the fumarate  
 RT reductase (complex II) of the anaerobic parasitic nematode, Ascaris  
 RT suum and the succinate dehydrogenase of the aerobic, free-living  
 RT nematode, Caenorhabditis elegans";  
 RL Mol. Biochem. Parasitol. 68:177-187(1994).  
 DR EMBL; D30650; BAA21636.1; -;  
 DR PFAM; PF00890; FAD\_binding\_2; 1.  
 DR PROSITE; PS00504; FRD\_SDH\_FAD\_BINDING; 1.  
 KW Oxidoreductase; Electron transport; Flavoprotein; FAD; Mitochondrion;  
 KW Transit peptide.  
 FT TRANSIT 1 30 MITOCHONDRION (POTENTIAL).  
 FT CHAIN 31 645 FLAVOPROTEIN SUBUNIT OF COMPLEX II.  
 SQ SEQUENCE 645 AA; 71200 MW; 3EBFA08D CRC32;

Query Match 83.8%; Score 31; DB 8; Length 645;  
 Best Local Similarity 85.7%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELRIKI 7  
 Db 525 KELRLK 531  
 |||||:|  
 |||||:|

RESULT 13  
 O29765 PRELIMINARY; PRT; 245 AA.  
 ID O29765  
 AC O29765;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
 DE HYPOTHETICAL 27.4 KD PROTEIN.  
 GN AF0485.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 OC Archaeoglobus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE; 98049343.  
 RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,  
 RA RICHARDSON D.L., KERLAVAGE A.J., GRAHAM D.E., KYRPIDES N.C.,  
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
 RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLOBE A., ZHOU L.,  
 RA OVERBEER K., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,  
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,  
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,  
 RA VENTER J.C.;  
 RT "The complete genome sequence of the hyperthermophilic,  
 RT sulphate-reducing archaeon Archaeoglobus fulgidus";  
 RL Nature 390:364-370(1997).  
 DR EMBL; AE001070; AAB90754.1; -;  
 DR TIGR; AF0485; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 245 AA; 27371 MW; D3BFF7CD CRC32;

Query Match 81.1%; Score 30; DB 1; Length 245;

RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 burgdorferi.";  
 RL Nature 390:580-586(1997).  
 DR EMBL; AE000785; AAC66050.1; -  
 DR TIGR; BBEL6; -  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 206 AA; 24534 MW; D085EC4D CRC32;

Query Match 83.8%; Score 31; DB 2; Length 206;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 |:|:|:|  
 Db 176 KKLRLKI 183

RESULT 7  
 O50986 PRELIMINARY; PRT; 309 AA.  
 ID O50986;  
 AC O50986;  
 DT 01-JUN-1998 (TREMELrel. 06, Created)  
 DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)  
 DE CONSERVED HYPOTHETICAL PROTEIN.  
 GN BBB10.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OG Plasmid cp26.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE; 98065943.

RA FRASER C.M., CASSENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTERBACK T., WATHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 burgdorferi.";  
 RL Nature 390:580-586(1997).  
 DR EMBL; AE000792; AAC66320.1; -  
 DR TIGR; BBB10; -  
 KW Plasmid.  
 SQ SEQUENCE 309 AA; 37272 MW; 877229A6 CRC32;

Query Match 83.8%; Score 31; DB 2; Length 309;  
 Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 |:|:|:|  
 Db 127 KQLAKIKI 134

RESULT 8  
 Q12009 PRELIMINARY; PRT; 286 AA.  
 ID Q12009;  
 AC Q12009;  
 DT 01-NOV-1996 (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMELrel. 01, Last annotation update)  
 DE ORF D1075.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomycetes.

RN SEQUENCE FROM N.A.  
 RP STRAIN-FY1679;  
 RA BAHR A., MOELLER-RIEKER S., HANKELN T., KRAEMER C., SCHMIDT E.R.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA SCHMIDT E.R., BAHR A., KRAEMER C., HANKELN T., MOELLER-RIEKER S.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X98000; CAA67468.1; -  
 DR EMBL; Z74249; CAA98779.1; -  
 SQ SEQUENCE 286 AA; 33391 MW; 1FF3675B CRC32;

Query Match 83.8%; Score 31; DB 3; Length 286;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 |:|:|:|  
 Db 26 KELKHVKI 33

RESULT 9  
 O74991 PRELIMINARY; PRT; 577 AA.  
 ID O74991;  
 AC O74991;  
 DT 01-NOV-1998 (TREMELrel. 08, Created)  
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
 DE HYPOTHETICAL 66.2 KD PROTEIN.  
 GN SPCC338.16.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H;  
 RA LYNE M., WOOD V., RAJANDREAM M.A., BARRELL B.G., MURPHY L.,  
 RA HARRIS D.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL023781; CAA19347.1; -  
 DR PFAM; PF00646; F-box; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 577 AA; 66238 MW; B84FCFF9 CRC32;

Query Match 83.8%; Score 31; DB 3; Length 577;  
 Best Local Similarity 87.5%; Pred. No. 3.2e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 |:|:|:|  
 Db 529 KELKRIDI 536

RESULT 10  
 Q93895 PRELIMINARY; PRT; 150 AA.  
 ID Q93895;  
 AC Q93895;  
 DT 01-FEB-1997 (TREMELrel. 02, Created)  
 DT 01-FEB-1997 (TREMELrel. 02, Last sequence update)  
 DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)  
 DE M03B6.1 PROTEIN.  
 GN M03B6.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL034559; CAB39033.1; -  
 SQ SEQUENCE 1844 AA; 221495 MW; C852E25E CRC32;

Query Match 86.5%; Score 32; DB 5; Length 1844;  
 Best Local Similarity 75.0%; Pred. No. 5.7e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KELRIKI 8  
 ||:|:|  
 Db 833 KEIKIKI 840

RESULT 3  
 ID Q58524 PRELIMINARY; PRT; 1195 AA.  
 AC Q58524;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE HYPOTHETICAL PROTEIN MJ1124.  
 GN MJ1124.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,  
 RA OVERBEK R., KIRKAGEN N.F., WEINSTOCK K.G., MERRICK J.M., GLOCKE A.,  
 RA SCOTT J.L., GEORHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST H.A., RAINE B.P., BORODOVSKI M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii."  
 RL Science 273:1058-1073(1996).  
 DR EMBL; U67555; AAB99126.1; -  
 DR PFAM; PF00270; DEAD; 1.  
 KW Hypothetical protein.  
 FT DOMAIN 137 142 POLY-VAL.  
 SQ SEQUENCE 1195 AA; 138673 MW; F03977D2 CRC32;

Query Match 83.8%; Score 31; DB 1; Length 1195;  
 Best Local Similarity 62.5%; Pred. No. 5.9e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KELRIKI 8  
 ||:|:|  
 Db 802 KEIKIKI 809

RESULT 4  
 ID O53029 PRELIMINARY; PRT; 169 AA.  
 AC O53029;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)  
 DE HYPOTHETICAL 19.1 KD PROTEIN.  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;  
 OC Enterococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S-48;  
 RX MEDLINE; 98143421.  
 RA MARTINEZ-BUENO M., VALDIVIA E., GLVEZ A., COYETTE J., MAQUEDA M.;  
 RT "Analysis of the gene cluster involved in production and immunity of

RT the peptide antibiotic AS-48 in Enterococcus faecalis."  
 RL Mol. Microbiol. 27:347-358(1998).  
 DR EMBL; Y12234; CAA72923.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 169 AA; 19070 MW; F9E19FB5 CRC32;

Query Match 83.8%; Score 31; DB 2; Length 169;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KELRIKI 8  
 ||:|:|  
 Db 6 KEVKIKI 13

RESULT 5  
 ID O52968 PRELIMINARY; PRT; 169 AA.  
 AC O52968;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE BACF PROTEIN.  
 GN BACF.  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OG Plasmid pPD1.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;  
 OC Enterococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98062218.  
 RA TOMITA H., FUJIMOTO S., TANIMOTO K., IKE Y.;  
 RT "Cloning and genetic and sequence analyses of the bacteriocin 21  
 determinant encoded on the Enterococcus faecalis pheromone-responsive  
 RT conjugative plasmid pPD1."  
 RL J. Bacteriol. 179:7843-7855(1997).  
 DR EMBL; D85752; BAA24810.1; -  
 KW Plasmid.  
 SQ SEQUENCE 169 AA; 19100 MW; 5BBC7025 CRC32;

Query Match 83.8%; Score 31; DB 2; Length 169;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KELRIKI 8  
 ||:|:|  
 Db 6 KEVKIKI 13

RESULT 6  
 ID O50712 PRELIMINARY; PRT; 206 AA.  
 AC O50712;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 24.5 KD PROTEIN.  
 GN BBE16.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OG Plasmid lp25.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:33 ; Search time 160.2 seconds  
(without alignments)  
3.073 Million cell updates/sec

Title: US-09-124-280A-25

Perfect score: 37

Sequence: 1 KELKRIKI 8

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- SPTREMBL\_10.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_prodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	36	97.3	440	1	O58231	O58231 Pyrococcus
2	32	86.5	1844	5	O97287	O97287 plasmodium
3	31	83.8	1195	1	O58524	O58524 methanococ
4	31	83.8	169	2	O53029	O53029 enterococc
5	31	83.8	169	2	O52968	O52968 enterococc
6	31	83.8	206	2	O50712	O50712 borrelia bu
7	31	83.8	309	2	O50986	O50986 borrelia bu
8	31	83.8	286	3	O12009	O12009 saccharomyc
9	31	83.8	577	3	O74991	O74991 schizosacch
10	31	83.8	150	5	O93895	O93895 caenorhabdi
11	31	83.8	980	5	O96246	O96246 plasmodium
12	31	83.8	645	8	O33862	O33862 ascaris suu
13	30	81.1	245	1	O29765	O29765 archaeoglob
14	30	81.1	67	1	O73967	O73967 pyrococcus
15	30	81.1	499	1	O58391	O58391 pyrococcus
16	30	81.1	788	1	O58603	O58603 pyrococcus
17	30	81.1	103	2	O25202	O25202 helicobacte
18	30	81.1	218	2	O25606	O25606 helicobacte
19	30	81.1	362	5	O18440	O18440 caenorhabdi
20	30	81.1	732	5	O24193	O24193 drosophila
21	30	81.1	541	5	O17848	O17848 caenorhabdi
22	30	81.1	225	5	O45913	O45913 caenorhabdi
23	30	81.1	1148	5	O77203	O77203 dictyosteli
24	30	81.1	1308	11	O922N7	O922N7 rattus norv
25	30	81.1	2165	12	O82021	O82021 human respi
26	30	81.1	2165	12	P90197	P90197 human respi
27	30	81.1	2165	12	O82027	O82027 human respi
28	30	81.1	2165	12	O41356	O41356 human respi
29	30	81.1	2165	12	O09721	O09721 respiratory

30	30	81.1	2165	12	O41355	O41355 human respi
31	30	81.1	1490	13	P79922	P79922 xenopus lae
32	29	78.4	361	1	O57936	O57936 pyrococcus
33	29	78.4	373	2	O66875	O66875 aquifex aeo
34	29	78.4	1156	2	O66878	O66878 aquifex aeo
35	29	78.4	241	2	O84728	O84728 chlamydia t
36	29	78.4	114	2	O57459	O57459 coxiella bu
37	29	78.4	208	2	O50819	O50819 borrelia bu
38	29	78.4	1008	2	O92D23	O92D23 rickettsia
39	29	78.4	235	2	O92742	O92742 chlamydia p
40	29	78.4	3262	4	O13788	O13788 homo sapien
41	29	78.4	858	4	O60346	O60346 homo sapien
42	29	78.4	812	4	O07902	O07902 homo sapien
43	29	78.4	1115	5	O43886	O43886 drosophila
44	29	78.4	1115	5	O44387	O44387 drosophila
45	29	78.4	1726	11	O62383	O62383 mus musculus

## ALIGNMENTS

### RESULT 1

O58231 ID O58231 PRELIMINARY; PRT; 440 AA.  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE 440AA LONG HYPOTHETICAL PROTEIN.  
 GN PH0495.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE; 98344137.  
 RA KAWABAYASHI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
 RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,  
 RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
 RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
 RA AKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
 RA KIKUCHI H.;  
 RT "Complete sequence and gene organization of the genome of a  
 hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998)  
 DR EMBL; AP000002; BAA29583.1; -  
 SQ SEQUENCE 440 AA; 473/74 MW; 776F69BE CRC32;

Query Match 97.3%; Score 36; DB 1; Length 440;  
 Best Local Similarity 87.5%; Pred. No. 31;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELKRIKI 8  
 DB 86 KELKRIKV 93  
 |||||

### RESULT 2

O97287 ID O97287 PRELIMINARY; PRT; 1844 AA.  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE MAL3P7.15. PROTEIN.  
 GN MAL3P7.15.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=307;  
 RA OLIVER K., BOWMAN S., HARRIS D., LAWSON D., QUAIL M., BARRELL B.;

```

RN  SEQUENCE FROM N.A.
RA  HEBLING U., HOFMANN B., DELIUS H.
RL  SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
RP  SEQUENCE FROM N.A.
RC  STRAIN-S288C;
RX  MEDLINE; 97435491.
RA  RIEGER M., BRUECKNER M., SCHAEFER M., MUELLER-AUER S.;
RT  "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RL  chromosome VII.";
CC  YEAST 13.1077-1090(1997).
CC  -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF PHOSPHOLIPID
CC  BIOSYNTHESIS AND INTERORGANELLE TRAFFICKING OF PHOSPHATIDYLSEINE.
CC  -1- CATALYTIC ACTIVITY: PHOSPHATIDYL-L-SERINE -
CC  PHOSPHATIDYLETHANOLAMINE + CO(2).
CC  -1- COFACTOR: THIS ENZYME REQUIRES A PYRUVYL GROUP FOR ITS ACTIVITY.
CC  -1- PATHWAY: AMINOPHOSPHOLIPID BIOSYNTHESIS.
CC  -1- SUBCELLULAR LOCATION: GOLGI STACK AND VACUOLES.
CC  -1- SIMILARITY: THE C-TERMINAL (FROM AA 788 ONWARD) IS SIMILAR TO
CC  C.PASTEURIANUM PSD-LIKE PROTEIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U19910; G841244; -.
DR  EMBL; 272955; E243354; -.
DR  SGD; L0002817; PSD2.
DR  PFAM; PF00168; C2; 1.
KW  PHOSPHOLIPID BIOSYNTHESIS; LYASE; DECARBOXYLASE; PYRUVATE; ZYMOGEN;
KW  GOLGI STACK.
FT  DOMAIN 152 159 POLY-SER.
FT  DOMAIN 271 274 POLY-SER.
FT  DOMAIN 591 594 POLY-GLN.
FT  SITE 1042 1043 CLEAVAGE (NONHYDROLYTICAL)
FT  (BY SIMILARITY).
FT  MOD_RES 1043 1043 CONVERTED TO A PYRUVYL GROUP
FT  (BY SIMILARITY).
FT  DOMAIN 1045 1049 POLY-ILE.
FT  CONFLICT 801 801 E -> G (IN REF. 1).
FT  CONFLICT 974 974 Y -> N (IN REF. 1).
SQ  SEQUENCE 1138 AA; 130064 MW; 63169611 CRC32;

Query Match 78.4%; Score 29; DB 1; Length 1138;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ELKRIKI 8
Db 1086 ELKRIKI 1092
|||||
ID MYSD_CHICK STANDARD; PRT; 1829 AA.
AC Q02440;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE DILUTE MYOSIN HEAVY CHAIN, ISOFORM I (MYOSIN HEAVY CHAIN P190)
DE (MYOSIN-V).
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANINAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;

```

```

RX MEDLINE; 93012002.
RA SANDERS G., LICHTER B., MEYER H.E., KILIMANN M.W.;
RT "CDNA encoding the chicken ortholog of the mouse dilute gene product.
RT Sequence comparison reveals a myosin I subfamily with conserved C-
RT terminal domains.";
RL FEBS LETT. 311:295-298(1992).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1154-1163.
RC TISSUE-BRAIN;
RX MEDLINE; 93107155.
RA ESPRAFICO E.M., CHENEY R.E., MATTEOLI M., NASCIMENTO A.A.,
RA DE CAMILLI P.V., LARSON R.E., MOOSEKER M.S.;
RT "Primary structure and cellular localization of chicken brain
RT myosin-V (p190), an unconventional myosin with calmodulin light
RT chains.";
RL J. CELL BIOL. 119:1541-1557(1992).
CC -1- SUBUNIT: MAY BE A HOMODIMER, WHICH ASSOCIATES WITH MULTIPLE
CC CALMODULIN OR MYOSIN LIGHT CHAINS.
CC -1- SUBCELLULAR LOCATION: GOLGI-DERIVED CYTOPLASMIC MEMBRANES
CC (POTENTIAL).
CC -1- TISSUE SPECIFICITY: NEURONAL AND NONNEURONAL CELLS OF THE BRAIN.
CC -1- SIMILARITY: BELONGS TO CLASS-5 MYOSINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; X67251; G63365; -.
DR  EMBL; Z11718; E45207; ALT_INIT.
DR  PFAM; PF00063; myosin_head; 1.
DR  PFAM; PF00612; IQ; 6.
DR  HSP; P08799; 1MND.
KW MYOSIN; REPEAT; ATP-BINDING; CALMODULIN-BINDING; ACTIN-BINDING;
KW HEPTAD REPEAT PATTERN; COILED COIL; PHOSPHORYLATION.
FT  DOMAIN 1 767 HEAD.
FT  DOMAIN 768 914 NECK.
FT  DOMAIN 915 1041 COILED COIL.
FT  DOMAIN 1042 1829 TAIL.
FT  NP_BIND 163 170 ATP (BY SIMILARITY).
FT  DOMAIN 644 666 ACTIN-BINDING (POTENTIAL).
FT  DOMAIN 774 788 CALMODULIN-BINDING (BY SIMILARITY).
FT  DOMAIN 797 811 CALMODULIN-BINDING (BY SIMILARITY).
FT  DOMAIN 822 836 CALMODULIN-BINDING (BY SIMILARITY).
FT  DOMAIN 845 863 CALMODULIN-BINDING (BY SIMILARITY).
FT  DOMAIN 870 885 CALMODULIN-BINDING (BY SIMILARITY).
FT  DOMAIN 893 907 CALMODULIN-BINDING (BY SIMILARITY).
FT  MOD_RES 1734 1734 PHOSPHORYLATION (POTENTIAL).
FT  CONFLICT 1142 1142 E -> EQ (IN REF. 2).
SQ  SEQUENCE 1829 AA; 212381 MW; 1F30669B CRC32;

```

Query Match 78.4%; Score 29; DB 1; Length 1829;  
 Best Local Similarity 62.5%; Pred. No. 4e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KELKRIKI 8  
 :|||:|  
 Db 904 RELKRLKI 911

Search completed: September 7, 1999, 23:59:06  
 Job time: 518 sec

polymorphism in the apolipoprotein B gene.;

[17] HUM. GENET. 86:91-93(1990).

RP VARIANT FDB GLN-3527.

RX MEDLINE; 89098975.

RA SORIA L.F., LUDWIG E.H., CLARKE H.R.G., VEGA G.L., GRUNDY S.M., MCCARTHY B.J.;

RA "Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100.;"

RT "Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100.;"

RL PROC. NATL. ACAD. SCI. U.S.A. 86:587-591(1989).

[18] VARIANT LEU-2739.

RP MEDLINE; 91016974.

RA HUANG L.-S., GAVISH D., BRESLOW J.L.;

RL "Sequence polymorphism in the human apob gene at position 8344.;"

RT NUCLEIC ACIDS RES. 18:5922-5922(1990).

[19] VARIANT FDB CYS-3558.

RX MEDLINE; 95190020.

RA PULLINGER C.R., HENNESSY L.K., CHATTERTON J.E., LIU W., LOVE J.A., MENDEL C.M., FROST P.H., MALLOY M.J., SCHUMAKER V.N., KANE J.P.;

RA "Familial ligand-defective apolipoprotein B. Identification of a new mutation that decreases LDL receptor binding affinity.;"

RT J. CLIN. INVEST. 95:1225-1234(1995).

[20] VARIANTS L-1437; S-1914; K-2566; T-3121; A-3945; M-4128 AND T-4481.

RP MEDLINE; 97044521.

RA POIRIER O., RICARD S., BEHAGUE I., SOURIAU C., EVANS A.E., ARVILLER D., MARQUES-VIDAL P., LUC G., ROIZES G., CAMBIEN F.;

RT "Detection of new variants in the apolipoprotein B (Apo B) gene by PCR-SSCP.;"

RL HUM. MUTAT. 8:282-285(1996).

[21] VARIANTS FDB GLN-3527 AND CYS-3558.

RX MEDLINE; 97403938.

RA RABES J.P., VARRET M., SAINT-JORE B., ERLICH D., JONDEAU G., KREMPF M., GIRAUDET P., JUNIEN C., BOILEAU C.;

RA "Familial ligand-defective apolipoprotein B-100: simultaneous detection of the ARG3500-->GLN and ARG3531-->CYS mutations in a French population.;"

RT HUM. MUTAT. 10:160-163(1997).

[22] VARIANTS S-1914; R-1923; L-2739; D-3319; T-3427; Q-3432 AND I-3921.

RP MEDLINE; 98141125.

RA LEREN T.P., BAKKEN K.S., HOEL V., HJERMANN I., BERG K.;

RL "Screening for mutations of the apolipoprotein B gene causing hypocholesterolemia.;"

RT HUM. GENET. 102:44-49(1998).

-1- FUNCTION: APOLOPROTEIN B IS A MAJOR PROTEIN CONSTITUENT OF CHYLOMICRONS, VLDL AND LDL. IT FUNCTIONS AS A RECOGNITION SIGNAL FOR THE CELLULAR BINDING AND INTERNALIZATION OF LDL PARTICLES BY THE APOB/E RECEPTOR.

-1- DISEASE: DEFECTS IN APOB ARE A CAUSE OF FAMILIAL LIGAND-DEFECTIVE APOLOPROTEIN B-100 (FDB). IT IS A DOMINANTLY INHERITED DISORDER OF LIPOPROTEIN METABOLISM LEADING TO HYPERCHOLESTEROLEMIA AND INCREASED PRONENESS TO CORONARY ARTERY DISEASE (CAD).

-1- IMPAIRED CLEARANCE OF LDL PARTICLES BY DEFECTIVE APOB/E RECEPTORS. (POLYGENIC) CAN BE THE CAUSE OF HYPOCHOLESTEROLEMIA.

-1- APO B-48 WHICH IS SYNTHESIZED ONLY BY THE INTESTINE, AND FOUND IN CHYLOMICRONS, IS A SHORTENED FORM OF APO B-100 WHICH LACKS THE LDL-RECEPTOR REGION. B-48 IS PRODUCED BY THE EDITING OF THE MRNA ON

Query Match 78.4%; Score 29; DB 1; Length 4563;  
Best Local Similarity 75.0%; Pred. No. 1e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELRIKI 8  
|:|:|:|

Db 2634 KDLNKIKI 2641

RESULT 13  
DCLY\_SALTY STANDARD; PRT; 715 AA.  
ID AC Q60002;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE LYSINE DECARBOXYLASE, INDUCIBLE (EC 4.1.1.18) (LDC).  
GN CADA.  
OS SALMONELLA TYPHIMURIUM.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE.  
OC SALMONELLA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96347132.  
RA PARK Y.K., BEARSON B., BANG S.H., BANG I.S., FOSTER J.W.;

RT "Internal pH crisis, lysine decarboxylase and the acid tolerance response of Salmonella typhimurium.;"

RL MOL. MICROBIOL. 20:605-611(1996).

CC -1- CATALYTIC ACTIVITY: L-LYSINE - CADAVERINE + CO(2).

CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).

CC -1- SIMILARITY: TO FAMILY 1 OF ORNITHINE, LYSINE, AND ARGININE DECARBOXYLASES.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL: U37109; G1046369; -  
DR STGENE; SG?????; CADA.  
DR PROSITE; PS00703; OKR\_DC\_1; 1.  
DR PFAM; PF01276; OKR\_DC\_1; 1.  
DR HSSP; P43099; 10RD.  
KW LYASE; DECARBOXYLASE; PYRIDOXAL PHOSPHATE; MULTIGENE FAMILY.  
FT BINDING 367 367 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
SQ SEQUENCE 715 AA; 81087 MW; 1A45770F CRC32;

Query Match 78.4%; Score 29; DB 1; Length 715;  
Best Local Similarity 71.4%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELRIK 7  
|:|:|:|

Db 437 KEIKRLK 443

RESULT 14  
DPS2\_YEAST STANDARD; PRT; 1138 AA.  
ID AC P30307;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE PHOSPHATIDYLSELINE DECARBOXYLASE PROENZYME 2 PRECURSOR (EC 4.1.1.65).  
GN PSD2 OR YGR170W.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-PT36;  
RX MEDLINE; 95197637.

RA TROTTER P.J., PEDRETTI J., YATES R., VOELKER D.R.;

RT "Phosphatidylserine decarboxylase 2 of Saccharomyces cerevisiae. Cloning and mapping of the gene, heterologous expression, and creation of the null allele.;"

RL J. BIOL. CHEM. 270:6071-6080(1995).



Matches 5: Conservative 3: Mismatches 0: Indels 0: Gaps 0:

Oy 1 KELRIKI 8  
 Db 445 QEMRIKV 452

RESULT 12  
 APB\_HUMAN  
 ID APB\_HUMAN STANDARD; PRT; 4563 AA.  
 AC P04114; Q00502;  
 DT 01-NOV-1986 (REL. 03, CREATED)  
 DT 01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE APOLIPOPROTEIN B-100 PRECURSOR (APO B-100) [CONTAINS: APOLIPOPROTEIN  
 B-48 (APO B-48)].  
 GN APOB.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 87016385.  
 RA KNOTT T.C., WALLIS S.C., POWELL L.M., PEASE R.J., LUSIS A.J.,  
 RA BLACKHART B., MCCARTHY B.J., MAHLEY R.W., LEVY-WILSON B., SCOTT J.;  
 RT "Complete cDNA and derived protein sequence of human apolipoprotein  
 B-100.";  
 RL NUCLEIC ACIDS RES. 14:7501-7503(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88003974.  
 RA LUDWIG E.H., BLACKHART B.D., PIROTTI V.R., CAIATI L., FORTIER C.,  
 RA KNOTT T., SCOTT J., MAHLEY R.W., LEVY-WILSON B., MCCARTHY B.J.;  
 RT "DNA sequence of the human apolipoprotein B gene.";  
 RL DNA 6:363-372(1987).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 87008488.  
 RA CHEN S.-H., YANG C.-Y., CHEN P.-F., SETZER D., TANIMURA M., LI W.-H.,  
 RA GOTTO A.M. JR., CHAN L.;  
 RT "The complete cDNA and amino acid sequence of human apolipoprotein  
 B-100.";  
 RL J. BIOL. CHEM. 261:12918-12921(1986).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 87041416.  
 RA LAW S.W., GRANT S.M., HIGUCHI K., HOSPATTANKAR A.V., LACKNER K.J.,  
 RA LEE N., BREWER H.B. JR.;  
 RT "Human liver apolipoprotein B-100 cDNA: complete nucleic acid and  
 RT derived amino acid sequence.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 83:8142-8146(1986).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 87161758.  
 RA CLADARAS C., HADZOPULOU-CLADARAS M., NOLTE R.T., ATKINSON D.,  
 RA ZANNIS V.I.;  
 RT "The complete sequence and structural analysis of human  
 RT apolipoprotein B-100: relationship between apoB-100 and apoB-48  
 RT forms.";  
 RL EMBO J. 5:3495-3507(1986).  
 RN [6]  
 RP SEQUENCE OF 709-906 FROM N.A.  
 RX MEDLINE; 85270450.  
 RA DEEB S.S., MOTULSKY A.G., ALBERS J.J.;  
 RT "A partial cDNA clone for human apolipoprotein B.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 82:4983-4986(1985).  
 RN [7]  
 RP SEQUENCE OF 3056-3159 FROM N.A.  
 RX MEDLINE; 86041888.  
 RA MEHRABIAN M., SCHUMAKER V.N., FAREED G.C., WEST R., JOHNSON D.F.,  
 RA KIRCHGESSNER T.G., LIN H.-C., WANG X., MA Y., MENDIAZ E., LUSIS A.J.;  
 RT "Human apolipoprotein B: identification of cDNA clones and  
 RT characterization of mRNA.";

RL NUCLEIC ACIDS RES. 13:6937-6953(1985).  
 RN [8]  
 RP SEQUENCE OF 1937-2018 AND 3811-4334 FROM N.A.  
 RX MEDLINE; 86093680.  
 RA CARLSSON P., OLOFSSON S.O., BONDJERS G., DARNFORS C., WIKLUND O.,  
 RA BJURSELL G.;  
 RT "Molecular cloning of human apolipoprotein B cDNA.";  
 RL NUCLEIC ACIDS RES. 13:8813-8826(1985).  
 RN [9]  
 RP SEQUENCE OF 3109-4563 FROM N.A.  
 RX MEDLINE; 85300528.  
 RA KNOTT T.J., RALL S.C. JR., INNERARITY T.L., JACOBSON S.F.,  
 RA URDEA M.J., LEVY-WILSON B., POWELL L.M., PEASE R.J., EDDY R.,  
 RA NAKAI H., BYERS M., PRIESTLEY L.M., ROBERTSON E., RALL L.B.,  
 RA BTHSHOLTZ C., SHOWS T.B., MAHLEY R.W., SCOTT J.;  
 RT "Human apolipoprotein B: structure of carboxyl-terminal domains,  
 RT sites of gene expression, and chromosomal localization.";  
 RL SCIENCE 230:137-43(1985).  
 RN [10]  
 RP SEQUENCE OF 1-291 FROM N.A.  
 RX MEDLINE; 86149325.  
 RA PROTTER A.A., HARDMAN D.A., SCHILLING J.W., MILLER J., APPLEBY V.,  
 RA CHEN G.C., KIRSHER S.W., MCENROE G., KANE J.P.;  
 RT "Isolation of a cDNA clone encoding the amino-terminal region of  
 RT human apolipoprotein B.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 83:1467-1471(1986).  
 RN [11]  
 RP SEQUENCE OF 1-1670 FROM N.A.  
 RX MEDLINE; 86287319.  
 RA PROTTER A.A., HARDMAN D.A., SATO K.Y., SCHILLING J.W.,  
 RA YAMANAKA M., HORT Y.J., HJERRILD K.A., CHEN G.C., KANE J.P.;  
 RT "Analysis of cDNA clones encoding the entire B-26 region of human  
 RT apolipoprotein B.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 83:5678-5682(1986).  
 RN [12]  
 RP PARTIAL SEQUENCE, AND IDENTIFICATION (APO-B48).  
 RX MEDLINE; 88018019.  
 RA CHEN S.-H., HABIB G., YANG C.-H., GU Z.-W., LEE B.R., WENG S.-H.,  
 RA SILBERMAN S.R., CAI S.-J., DESLYPERE J.P., ROSSENEU M.,  
 RA GOTTO A.M. JR., LI W.-H., CHAN L.;  
 RT "Apolipoprotein B-48 is the product of a messenger RNA with an organ-  
 RT specific in-frame stop codon.";  
 RL SCIENCE 238:363-366(1987).  
 RN [13]  
 RP DOMAINS.  
 RX MEDLINE; 87039351.  
 RA KNOTT T.C., PEASE R.J., POWELL L.M., WALLIS S.C., RALL S.C. JR.,  
 RA INNERARITY T.L., BLACKHART B., TAYLOR W.R., MARCEL Y., MILNE R.,  
 RA JOHNSON D., FULLER M., LUSIS A.J., MCCARTHY B.J., MAHLEY R.W.,  
 RA LEVY-WILSON B., SCOTT J.;  
 RT "Complete protein sequence and identification of structural domains  
 RT of human apolipoprotein B.";  
 RL NATURE 323:734-738(1986).  
 RN [14]  
 RP DOMAINS.  
 RX MEDLINE; 86242245.  
 RA TANIMURA M., LI W.-H., SPARROW D.A., DELOOF H., ROSSENEU M.,  
 RA LEE F.-S., GU Z.-W., GOTTO A.M. JR., CHAN L.;  
 RT "Sequence, structure, receptor-binding domains and internal repeats  
 RT of human apolipoprotein B-100.";  
 RL NATURE 323:738-742(1986).  
 RN [15]  
 RP CALCIUM-BINDING DATA.  
 RX MEDLINE; 86242245.  
 RA DASHITI N., LEE D.M., MOK T.;  
 RT "Apolipoprotein B is a calcium binding protein.";  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 137:493-499(1986).  
 RN [16]  
 RP VARIANT SER-4338.  
 RX MEDLINE; 91071750.  
 RA NAYARAS M., LAURENT A.-M., MOREEL J.-F., RAGAB A., CAMBOU J.-P.,  
 RA CUNNY G., CAMBIEN F., ROIZES G.;  
 RT "Detection by denaturing gradient gel electrophoresis of a new

SQ SEQUENCE 439 AA; 51855 MW; F7827EF7 CRC32;

Query Match 81.1%; Score 30; DB 1; Length 439;  
Best Local Similarity 85.7%; Pred. No. 61;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELARIK 7  
:|||||  
DB 282 RELARIK 288

RESULT 9  
RRPL\_HRSVA ID RP\_L\_HRSVA STANDARD; PRT; 2165 AA.  
AC P28887; 1992 (REL. 24, CREATED)  
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE RNA POLYMERASE BETA SUBUNIT (EC 2.7.7.48) (LARGE STRUCTURAL PROTEIN)  
GN L  
OS HUMAN RESPIRATORY SYNCYTIAL VIRUS (STRAIN A2).  
OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES;  
OC PARAMYXOVIRIDAE; PNEUMOVIRINAE; PNEUMOVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 91272488.  
RA STEC D.S., HILL M.G. III, COLLINS P.L.;  
RT "Sequence analysis of the polymerase L gene of human respiratory  
RT syncytial virus and predicted phylogeny of nonsegmented  
RT negative-strand viruses."  
RL VIROLOGY 183:273-287(1991).  
CC -!- FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY  
CC SYNTHESIS OF NEWLY SYNTHESIZED VIRAL MENAS, RNA EDITING OF THE P  
CC GENE TRANSCRIPT, AND PROTEIN KINASE ACTIVITY.  
CC -!- SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M75730; G333956;  
DR PIR: A40317; RNZA2.  
DR PFAM: PF00946; Paramyx\_RNA\_pol; 1.  
KW POLYMERASE; RNA-DIRECTED RNA POLYMERASE.  
SQ SEQUENCE 2165 AA; 250384 MW; 72CE0E98 CRC32;

Query Match 81.1%; Score 30; DB 1; Length 2165;  
Best Local Similarity 75.0%; Pred. No. 3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KELARIK 8  
:|||||  
DB 256 KELARIIV 263

RESULT 10  
YMI3\_YEAST ID YMI3\_YEAST STANDARD; PRT; 1267 AA.  
AC Q04217;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PUTATIVE ATP-DEPENDENT RNA HELICASE YMR128W.  
GN YMR128W OR YMR553.04.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

CC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
CC SACCHAROMYCETACEAE; SACCHAROMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA BADCOCK K., CHURCHER C., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;  
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.  
CC -!- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES. BELONGS TO THE  
CC "DEAH" SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: Z48622; G728667;  
DR PROSITE: PS00690; DEAH\_ATP\_HELICASE; 1.  
DR PFAM: PF00271; Helicase\_C; 1.  
KW HYPOTHETICAL PROTEIN; ATP-BINDING; RNA-BINDING; HELICASE.  
FT NP\_BIND 414 421 ATP (POTENTIAL).  
FT SITE 516 519 DEAH BOX.  
SQ SEQUENCE 1267 AA; 144954 MW; 0DB79D64 CRC32;

Query Match 81.1%; Score 30; DB 1; Length 1267;  
Best Local Similarity 85.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KELARIK 7  
:|||||  
DB 21 KELKRI 27

RESULT 11  
YPO4\_CAEEL ID YPO4\_CAEEL STANDARD; PRT; 1131 AA.  
AC P53015;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 127.2 KD PROTEIN F31E3.4 IN CHROMOSOME III.  
GN F31E3.4  
OS CAENORHABDITIS ELEGANS.  
CC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
CC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA FAVELLO A.;  
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: TO YEAST PAN2 AND TO S.POMBE SPAC22G7.04.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U40935; G1072160;  
DR WORMPEP: F31E3.4; CE01269.  
DR PFAM: PF00929; Exonuclease; 1.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 1131 AA; 127174 MW; FF0E5E02 CRC32;

Query Match 81.1%; Score 30; DB 1; Length 1131;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;

Query Match	Best Local Similarity	100.08;	Score 58;	DB 1;	Length 10;	Mismatches	0;	Indels	0;	Gaps	0;
1 KTKCKFLKCC 10	1 KTKCKFLKCC 10	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
2 KTKCKFLKCC 11	2 KTKCKFLKCC 11	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
<p>US-08-280-397-2</p> <p>Sequence 2, Application US/08280397</p> <p>Patent No. 5589459</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Porro, Massimo</p> <p>TITLE OF INVENTION: Synthetic Peptides for Detoxification</p> <p>TITLE OF INVENTION: of Bacterial Endotoxins and for the</p> <p>TITLE OF INVENTION: Prevention and Treatment of Septic</p> <p>TITLE OF INVENTION: Shock</p> <p>NUMBER OF SEQUENCES: 10</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Hedman, Gibson &amp; Costigan, P.C.</p> <p>STREET: 1185 Avenue of the Americas</p> <p>CITY: New York</p> <p>STATE: New York</p> <p>COUNTRY: USA</p> <p>ZIP: 10036</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage</p> <p>COMPUTER: IBM PS/2</p> <p>OPERATING SYSTEM: DOS</p> <p>SOFTWARE: Word Perfect 5.1</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/280.397</p> <p>FILING DATE: 07/26/94</p> <p>CLASSIFICATION: 514</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: 07/819,893</p> <p>FILING DATE: 01/16/92</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Costigan, James V.</p> <p>REGISTRATION NUMBER: 25,669</p> <p>REFERENCE/DOCKET NUMBER: 576-002A</p> <p>TELEPHONE: (212) 302-8989</p> <p>TELEFAX: (212) 302-8998</p> <p>INFORMATION FOR SEQ ID NO: 6:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 10 amino acids</p> <p>TYPE: amino acid</p> <p>TOPOLOGY: circular</p> <p>US-08-280-397-2</p>											
1 KTKCKFLKCC 10	1 KTKCKFLKCC 10	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
2 KTKCKFLKCC 11	2 KTKCKFLKCC 11	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
<p>US-08-280-397-2</p> <p>Sequence 6, Application US/08280397</p> <p>Patent No. 5589459</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Porro, Massimo</p> <p>TITLE OF INVENTION: Synthetic Peptides for Detoxification</p> <p>TITLE OF INVENTION: of Bacterial Endotoxins and for the</p> <p>TITLE OF INVENTION: Prevention and Treatment of Septic</p> <p>TITLE OF INVENTION: Shock</p> <p>NUMBER OF SEQUENCES: 10</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Hedman, Gibson &amp; Costigan, P.C.</p> <p>STREET: 1185 Avenue of the Americas</p> <p>CITY: New York</p> <p>STATE: New York</p> <p>COUNTRY: USA</p> <p>ZIP: 10036</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage</p> <p>COMPUTER: IBM PS/2</p> <p>OPERATING SYSTEM: DOS</p> <p>SOFTWARE: Word Perfect 5.1</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/280.397</p> <p>FILING DATE: 07/26/94</p> <p>CLASSIFICATION: 514</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: 07/819,893</p> <p>FILING DATE: 01/16/92</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Costigan, James V.</p> <p>REGISTRATION NUMBER: 25,669</p> <p>REFERENCE/DOCKET NUMBER: 576-002A</p> <p>TELEPHONE: (212) 302-8989</p> <p>TELEFAX: (212) 302-8998</p> <p>INFORMATION FOR SEQ ID NO: 2:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 10 amino acids</p> <p>TYPE: amino acid</p> <p>TOPOLOGY: circular</p> <p>US-08-280-397-2</p>											
1 KTKCKFLKCC 10	1 KTKCKFLKCC 10	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
2 KTKCKFLKCC 11	2 KTKCKFLKCC 11	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
<p>US-08-280-397-2</p> <p>Sequence 6, Application US/08280397</p> <p>Patent No. 5589459</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Porro, Massimo</p> <p>TITLE OF INVENTION: Synthetic Peptides for Detoxification</p> <p>TITLE OF INVENTION: of Bacterial Endotoxins and for the</p> <p>TITLE OF INVENTION: Prevention and Treatment of Septic</p> <p>TITLE OF INVENTION: Shock</p> <p>NUMBER OF SEQUENCES: 10</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Hedman, Gibson &amp; Costigan, P.C.</p> <p>STREET: 1185 Avenue of the Americas</p> <p>CITY: New York</p> <p>STATE: New York</p> <p>COUNTRY: USA</p> <p>ZIP: 10036</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage</p> <p>COMPUTER: IBM PS/2</p> <p>OPERATING SYSTEM: DOS</p> <p>SOFTWARE: Word Perfect 5.1</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/280.397</p> <p>FILING DATE: 07/26/94</p> <p>CLASSIFICATION: 514</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: 07/819,893</p> <p>FILING DATE: 01/16/92</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Costigan, James V.</p> <p>REGISTRATION NUMBER: 25,669</p> <p>REFERENCE/DOCKET NUMBER: 576-002A</p> <p>TELEPHONE: (212) 302-8989</p> <p>TELEFAX: (212) 302-8998</p> <p>INFORMATION FOR SEQ ID NO: 2:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 10 amino acids</p> <p>TYPE: amino acid</p> <p>TOPOLOGY: circular</p> <p>US-08-280-397-2</p>											
1 KTKCKFLKCC 10	1 KTKCKFLKCC 10	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
2 KTKCKFLKCC 11	2 KTKCKFLKCC 11	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
<p>US-08-280-397-2</p> <p>Sequence 6, Application US/08280397</p> <p>Patent No. 5589459</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Porro, Massimo</p> <p>TITLE OF INVENTION: Synthetic Peptides for Detoxification</p> <p>TITLE OF INVENTION: of Bacterial Endotoxins and for the</p> <p>TITLE OF INVENTION: Prevention and Treatment of Septic</p> <p>TITLE OF INVENTION: Shock</p> <p>NUMBER OF SEQUENCES: 10</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Hedman, Gibson &amp; Costigan, P.C.</p> <p>STREET: 1185 Avenue of the Americas</p> <p>CITY: New York</p> <p>STATE: New York</p> <p>COUNTRY: USA</p> <p>ZIP: 10036</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage</p> <p>COMPUTER: IBM PS/2</p> <p>OPERATING SYSTEM: DOS</p> <p>SOFTWARE: Word Perfect 5.1</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/280.397</p> <p>FILING DATE: 07/26/94</p> <p>CLASSIFICATION: 514</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: 07/819,893</p> <p>FILING DATE: 01/16/92</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Costigan, James V.</p> <p>REGISTRATION NUMBER: 25,669</p> <p>REFERENCE/DOCKET NUMBER: 576-002A</p> <p>TELEPHONE: (212) 302-8989</p> <p>TELEFAX: (212) 302-8998</p> <p>INFORMATION FOR SEQ ID NO: 2:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 10 amino acids</p> <p>TYPE: amino acid</p> <p>TOPOLOGY: circular</p> <p>US-08-280-397-2</p>											
1 KTKCKFLKCC 10	1 KTKCKFLKCC 10	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
2 KTKCKFLKCC 11	2 KTKCKFLKCC 11	100.08;	Score 58;	DB 1;	Length 10;	0;	0;	0;	0;	0;	0;
<p>US-08-280-397-2</p> <p>Sequence 6, Application US/08280397</p> <p>Patent No. 5589459</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Porro, Massimo</p> <p>TITLE OF INVENTION: Synthetic Peptides for Detoxification</p> <p>TITLE OF INVENTION: of Bacterial End</p>											

REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 600.286US01  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-218-026-49

Query Match 100.0%; Score 58; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.003;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
| | | | | | | | | |  
Db 1 KTKCKFLKKC 10

RESULT 8  
US-08-218-026-50  
Sequence 50, Application US/08218026  
Patent No. 5786324  
GENERAL INFORMATION:  
APPLICANT: Gray, Beulah  
APPLICANT: Haseman, Judith R.  
TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
TITLE OF INVENTION: Activity and Endotoxin Neutralizing Activity for Gram  
TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5786324west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/218,026  
FILING DATE: 24-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalchuk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 600.286US01  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
MOLECULE TYPE: peptide  
US-08-218-026-50

Query Match 100.0%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
| | | | | | | | | |

Db 2 KTKCKFLKKC 11  
RESULT 9  
US-08-653-632-49  
Sequence 49, Application US/08653632  
Patent No. 5830860  
GENERAL INFORMATION:  
APPLICANT: GRAY, Beulah  
APPLICANT: HASEMAN, Judith R.  
APPLICANT: MAYO, Kevin  
TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 5830860west Center, 90 South Seventh St  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/653,632  
FILING DATE: 24-MAY-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/218026  
FILING DATE: 24-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Kettelberger, Denise M  
REGISTRATION NUMBER: 33,924  
REFERENCE/DOCKET NUMBER: 600.286US11  
TELEPHONE: 612/371-5268  
TELEFAX: 612/332-9081  
TELEX:  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
US-08-653-632-49

Query Match 100.0%; Score 58; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.003;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
| | | | | | | | | |  
Db 1 KTKCKFLKKC 10

RESULT 10  
US-08-653-632-50  
Sequence 50, Application US/08653632  
Patent No. 5830860  
GENERAL INFORMATION:  
APPLICANT: GRAY, Beulah  
APPLICANT: HASEMAN, Judith R.  
APPLICANT: MAYO, Kevin  
TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
NUMBER OF SEQUENCES: 66

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
;; STREET: 3100 No. 5830860west Center, 90 South Seventh St  
;; CITY: Minneapolis  
;; STATE: MN  
;; COUNTRY: USA  
;; ZIP: 55402  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSEQ Version 1.5  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/653,632  
;; FILING DATE: 24-MAY-1996  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/218026  
;; FILING DATE: 24-MAR-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kettelberger, Denise M  
;; REGISTRATION NUMBER: 33,924  
;; REFERENCE/DOCKET NUMBER: 600.286US11  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 612/371-5268  
;; TELEFAX: 612/332-9081  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 50:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; ORIGINAL SOURCE:  
;; US-08-653-632-50

Query Match 100.0%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
Db 2 KTKCKFLKKC 11

RESULT 11  
US-08-456-112B-31  
;; Sequence 31, Application US/08456112B  
;; Patent No. 5834430  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995

;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 31:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; US-08-456-112B-31  
Query Match 100.0%; Score 58; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.003;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
Db 1 KTKCKFLKKC 10

RESULT 12  
US-08-456-112B-35  
;; Sequence 35, Application US/08456112B  
;; Patent No. 5834430  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 35:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; US-08-456-112B-35

Query Match 100.0%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTKCKFLKKC 10  
| | | | |  
Db 2 KTKCKFLKKC 11

## RESULT 13

US-08-049-871-1  
; Sequence 1, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-1

Query Match 74.1%; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CKFLKKC 10  
| | | | |  
Db 1 CKFLKKC 7

## RESULT 14

US-07-819-893-1  
; Sequence 1, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare

; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
US-07-819-893-1

Query Match 74.1%; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CKFLKKC 10  
| | | | |  
Db 1 CKFLKKC 7

## RESULT 15

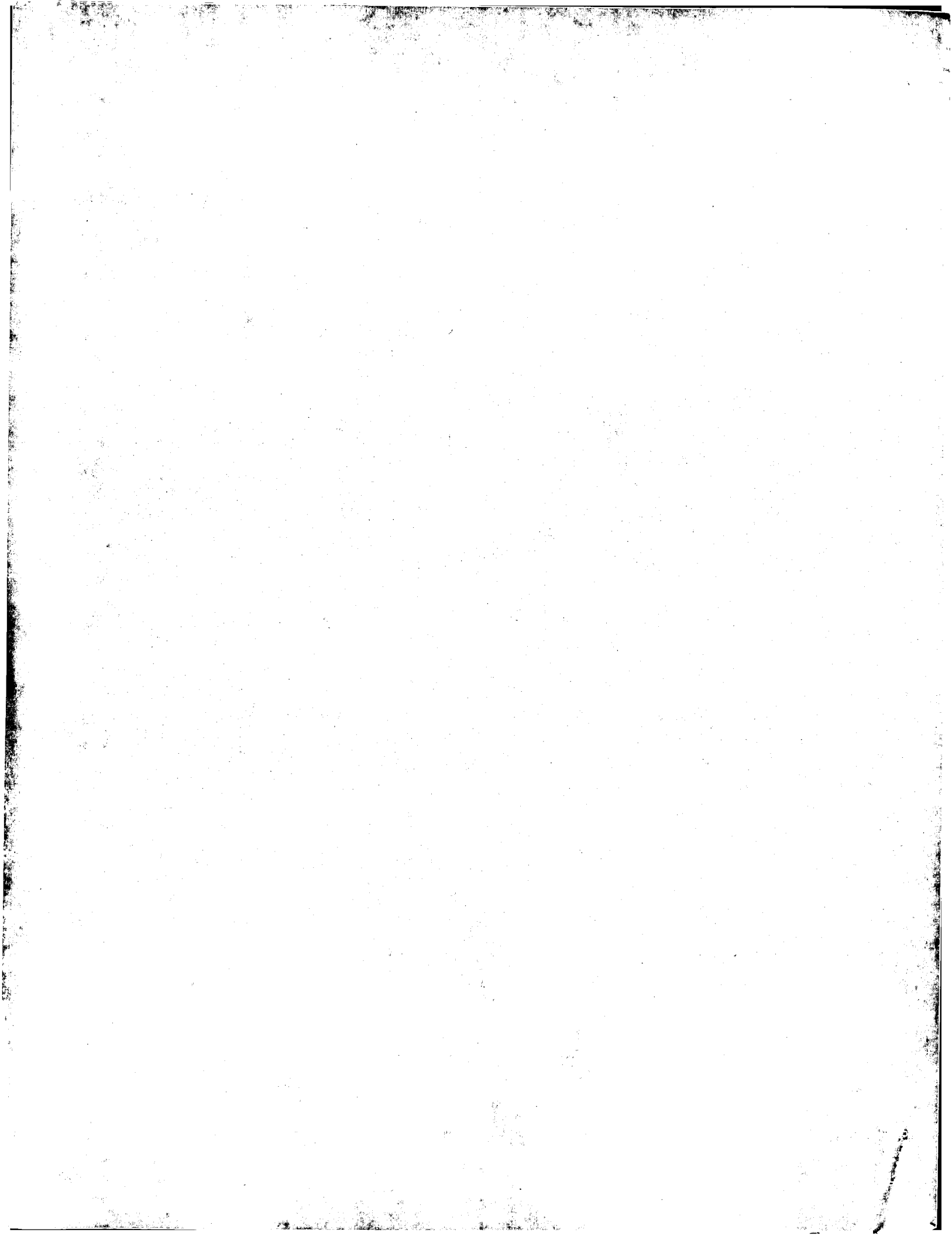
US-08-280-397-1  
; Sequence 1, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.

;  
;     REGISTRATION NUMBER: 25,669  
;     REFERENCE/DOCKET NUMBER: 576-002A  
;     TELECOMMUNICATION INFORMATION:  
;     TELEPHONE: (212) 302-8989  
;     TELEFAX: (212) 302-8998  
;     INFORMATION FOR SEQ ID NO: 1:  
;     SEQUENCE CHARACTERISTICS:  
;     LENGTH: 7 amino acids  
;     TYPE: amino acid  
;     TOPOLOGY: circular  
;     US-08-280-397-1

Query Match     74.18; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. NO. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY     4 CKFLKKC 10  
       | | | | |  
Db     1 CKFLKKC 7

Search completed: September 7, 1999, 23:07:49  
Job time: 1747 sec





GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:45 ; Search time 105.14 Seconds  
(without alignments)  
3.811 Million cell updates/sec

Title: US-09-124-280A-31  
Perfect score: 58  
Sequence: 1 KTKCKFLKCC 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	67.2	73	2 C44088	homeotic protein H
2	38.5	66.4	239	2 T03078	conserved hypothet
3	37	63.8	554	2 JU0188	membrane protein 4
4	36	62.1	445	2 G01452	NIMA-like protein
5	36	62.1	244	2 A49202	cysteine-rich secr
6	36	62.1	347	2 A70173	femA protein (femA
7	35	60.3	2749	1 ACMSIT	inositol 1,4,5-tri
8	35	60.3	2734	2 B36579	inositol 1,4,5-tri
9	35	60.3	2749	2 A36579	inositol 1,4,5-tri
10	35	60.3	2713	2 A55713	inositol 1,4,5-tri
11	35	60.3	2695	2 S54974	type 1 inositol 1,
12	35	60.3	1059	2 B24872	retrovirus-related
13	35	60.3	1040	2 S34639	pol protein - fru
14	35	60.3	135	2 C70028	conserved hypothet
15	34	58.6	1089	2 S33727	platelet-derived g
16	34	58.6	553	2 S16809	phosphoprotein pho
17	34	58.6	298	2 D69023	cell division inh
18	34	58.6	246	2 A64326	hypothetical prote
19	34	58.6	60	2 J50583	hypothetical 7k pr
20	34	58.6	206	2 S49882	hypothetical prote
21	34	58.6	1121	2 S54504	hypothetical prote
22	34	58.6	984	2 S14382	hypothetical prote
23	34	58.6	403	2 S07825	hypothetical prote
24	34	58.6	603	2 T00379	KIAA0640 protein -
25	34	58.6	189	2 C70485	hypothetical prote
26	33	56.9	223	1 Q0BE26	B2LF2 protein - hu
27	33	56.9	1003	1 P1VZAM	spheroidin precurs
28	33	56.9	547	2 A23595	DNA nucleotidyllexo
29	33	56.9	191	2 I45716	Cdc42 - fruit fly
30	33	56.9	3344	2 J01899	genome polypotein
31	33	56.9	491	2 B71957	ATP-dependent RNA
32	33	56.9	492	2 G64550	ATP-dependent RNA
33	33	56.9	284	2 S31424	DNA-binding homeot
34	33	56.9	284	2 A46304	homeotic protein H
35	33	56.9	282	2 B71444	hypothetical prote
36	33	56.9	355	2 E54445	formylmethanofuran
37	33	56.9	1766	2 S03701	141K protein - pea
38	33	56.9	318	2 T00402	homeodomain-zipper
39	33	56.9	301	2 S60936	LEE1 protein - yea

40 33 56.9 2380 2 E71604  
41 33 56.9 1122 2 F71606  
42 33 56.9 220 2 S29195  
43 33 56.9 311 2 S08110  
44 33 56.9 88 2 S61552  
45 33 56.9 760 2 I48745

ALIGNMENTS

RESULT 1  
C44088  
homeotic protein HAT22 - Arabidopsis thaliana (fragments)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 02-Jul-1998  
C:Accession: C44088  
R:Schen, M.; Davis, R.W.  
Proc. Natl. Acad. Sci. U.S.A. 89, 3894-3898, 1992  
A:Title: HD-zip proteins: members of an Arabidopsis homeodomain protein superfamily.  
A:Reference number: A44088; MUID:92237275  
A:Accession: C44088  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-73 <SCH>  
A:Cross-references: GB:M90417  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
F:18-32/Domain: homeobox homology (fragment) <HOX>

Query Match 67.2%; Score 39; DB 2; Length 73;  
Best Local Similarity 85.7%; Pred. No. 3.7;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CKFLKCC 10  
Db 39 CEFLKCC 45  
RESULT 2  
T03078  
conserved hypothetical protein 061L - Chilo iridescent virus  
C:Species: Chilo iridescent virus  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C:Accession: T03078  
R:Bahr, U.; Tidona, C.A.; Darai, G.  
Virus Genes 15, 235-245, 1997  
A:Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.  
A:Reference number: Z14834  
A:Accession: T03078  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-239 <BAH>  
A:Cross-references: EMBL:AF005334; NID:g2738385; PID:g2738425

Query Match 66.4%; Score 38.5; DB 2; Length 239;  
Best Local Similarity 63.6%; Pred. No. 11;  
Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;  
QY 1 KTK-CKFLKCC 10  
Db 83 KTRCKYVKKC 93  
RESULT 3  
JU0188  
membrane protein 4.1 homolog - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Feb-1994 #sequence\_revision 20-Aug-1994 #text\_change 21-Nov-1998  
C:Accession: I55505; JU0188  
R:Takeuchi, K.; Kawashima, A.; Nagafuchi, A.; Tsukita, S.  
J. Cell Sci. 107, 1921-1928, 1994

hypothetical prote  
probable multiple  
antistatin - Hydra  
salivary protein 1  
mammary tumor prot  
semaphorin B - mou

A>Title: Structural diversity of band4.1 superfamily members.  
 A:Reference number: 135505; MUID:95074267  
 A:Accession: 135505  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-554 <RES>  
 A:Cross-references: GB:D28818; NID:g466547; PID:g466548  
 C:Superfamily: protein 4.1 membrane-binding domain homology  
 F:13-295/Domain: protein 4.1 membrane-binding domain homology <B41>

Query Match 63.8%; Score 37; DB 2; Length 554;  
 Best Local Similarity 70.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10  
 || || || ||  
 Db 278 KTKCKLWKC 287

RESULT 4  
 G01452  
 NIMA-like protein kinase 1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 17-Jul-1998  
 C:Accession: G01452  
 R:Lu, K.P.  
 submitted to the EMBL Data Library, June 1994

A:Reference number: G07172  
 A:Accession: G01452  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-445 <LUX>  
 A:Cross-references: EMBL:U11050; NID:g507874; PID:g507875  
 C:Genetics:  
 A:Gene: NLK1  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology  
 F:6-271/Domain: protein kinase homology <KIN>  
 F:14-22/Region: protein kinase ATP-binding motif

Query Match 62.1%; Score 36; DB 2; Length 445;  
 Best Local Similarity 77.8%; Pred. No. 48;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTKCKFLKK 9  
 | ||| |||  
 Db 407 KSKCKDLKK 415

RESULT 5  
 A49202  
 Cysteine-rich secretory protein-1 - mouse  
 N:Alternate names: CRISP-1  
 C:Species: Mus musculus (house mouse)  
 C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Sep-1998  
 C:Accession: A49202  
 R:Haendler, B.; Kratzschmar, J.; Theuring, F.; Schleuning, W.D.  
 Endocrinology 133, 192-198, 1993  
 A>Title: Transcripts for cysteine-rich secretory protein-1 (CRISP-1; DE/ABG) and the novel  
 A:Reference number: A49202; MUID:93307144  
 A:Accession: A49202  
 A>Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-244 <HAE>  
 A:Cross-references: GB:I05559; NID:g309190; PID:g309191  
 A:Experimental source: NWRI, epididymis, salivary gland  
 A>Note: sequence extracted from NCBI backbone (NCBIN:134675, NCBI:134676).  
 C:Superfamily: cysteine-rich secretory protein 1

Query Match 62.1%; Score 36; DB 2; Length 244;  
 Best Local Similarity 75.0%; Pred. No. 30;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 TKCKFLKK 9  
 | ||| |||  
 Db 213 TNCKYLKK 220

RESULT 6  
 A70173  
 femA protein (femA) homolog - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
 C:Accession: A70173  
 R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A>Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: A70173  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-347 <XLE>  
 A:Cross-references: GB:AE001160; GB:AE000783; NID:g2688505; PID:g2688511; TIGR:BB0586  
 A:Experimental source: strain B31

Query Match 62.1%; Score 36; DB 2; Length 347;  
 Best Local Similarity 77.8%; Pred. No. 39;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTKCKFLKK 9  
 ||| | |||  
 Db 121 KTKIKLKK 129

RESULT 7  
 ACMSIT  
 inositol 1,4,5-trisphosphate receptor - mouse  
 N:Alternate names: inositol-1,4,5-trisphosphate-binding protein P400; membrane-associ  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 02-Jul-1998  
 C:Accession: S04844; S06796; C34955  
 R:Furuichi, T.; Yoshikawa, S.; Mikoshiba, K.  
 Nucleic Acids Res. 17, 5385-5386, 1989  
 A>Title: Nucleotide sequence of cDNA encoding P400 protein in the mouse cerebellum.  
 A:Reference number: S04844; MUID:89345101  
 A:Accession: S04844  
 A>Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-2749 <FUR>  
 A:Cross-references: EMBL:X15373; NID:g53568; PID:g53569  
 A:Experimental source: strain ICR; cerebellum  
 R:Furuichi, T.; Yoshikawa, S.; Miyawaki, A.; Wada, K.; Maeda, N.; Mikoshiba, K.  
 Nature 342, 32-38, 1989  
 A>Title: Primary structure and functional expression of the inositol 1,4,5-trisphosph  
 A:Reference number: S06796; MUID:90044039  
 A:Accession: S06796  
 A>Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-2749 <FU2>  
 A:Cross-references: EMBL:X15373; NID:g53568; PID:g53569  
 A:Experimental source: strain ICR; cerebellum  
 R:Nordquist, D.T.; Kozak, C.A.; Orr, H.T.  
 J. Neurosci. 8, 4780-4789, 1988  
 A>Title: cDNA cloning and characterization of three genes uniquely expressed in cereb  
 A:Reference number: A92975; MUID:89068131  
 A:Accession: C34955  
 A:Molecule type: mRNA  
 A:Residues: 2250-2674, 'p', 2676-2749 <NOR>  
 C:Superfamily: inositol-trisphosphate receptor  
 C:Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein

```
F:890-907/Domain: transmembrane #status predicted <TM1>
F:1960-1976/Domain: transmembrane #status predicted <TM2>
F:2276-2294/Domain: transmembrane #status predicted <TM3>
F:2308-2326/Domain: transmembrane #status predicted <TM4>
F:2334-2350/Domain: transmembrane #status predicted <TM5>
F:2352-2372/Domain: transmembrane #status predicted <TM6>
F:2391-2407/Domain: transmembrane #status predicted <TM7>
F:2440-2462/Domain: transmembrane #status predicted <TM8>
F:2570-2589/Domain: transmembrane #status predicted <TM9>
F:1588,1735/Binding site: phosphate (Ser) (covalent) #status predicted
F:2475,2503,2622,2710/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          60.3%; Score 35; DB 1; Length 2749;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
      | : ||: |||
Db 1317 KAEGKFIKCC 1326

RESULT 8
B36579
inositol 1,4,5-trisphosphate receptor 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 08-Sep-1997
C:Accession: B36579
R:Mignery, G.A.; Newton, C.L.; Archer III, B.T.; Suedhof, T.C.
J. Biol. Chem. 265, 12679-12685, 1990
A:Title: Structure and expression of the rat inositol 1,4,5-trisphosphate receptor.
A:Reference number: A36579; MUID:90324264
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2734 <MIG>
A:Cross-references: GB:J05510; NID:g204673; PID:g204674
C:Superfamily: inositol-trisphosphate receptor

Query Match          60.3%; Score 35; DB 2; Length 2734;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
      | : ||: |||
Db 1302 KAEGKFIKCC 1311

RESULT 9
A36579
inositol 1,4,5-trisphosphate receptor 1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 08-Sep-1997
C:Accession: A36579
R:Mignery, G.A.; Newton, C.L.; Archer III, B.T.; Suedhof, T.C.
J. Biol. Chem. 265, 12679-12685, 1990
A:Title: Structure and expression of the rat inositol 1,4,5-trisphosphate receptor.
A:Reference number: A36579; MUID:90324264
A:Accession: A36579
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2749 <MIG>
A:Cross-references: GB:J05510; NID:g204673; PID:g204675
C:Superfamily: inositol-trisphosphate receptor
C:Keywords: phosphoprotein; transmembrane protein

Query Match          60.3%; Score 35; DB 2; Length 2749;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
```

```
      | : ||: |||
Db 1317 KAEGKFIKCC 1326

RESULT 10
A55713
inositol 1,4,5-trisphosphate receptor type 1 - human
C:Species: Homo sapiens (man)
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 24-Sep-1998
C:Accession: A55713
R:Harnick, D.J.; Jayaraman, T.; Ma, Y.; Mullerl, P.; Go, L.O.; Marks, A.R.
J. Biol. Chem. 270, 2833-2840, 1995
A:Title: The human type 1 inositol 1,4,5-trisphosphate receptor from T lymphocytes. S
A:Reference number: A55713; MUID:95155356
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2713 <HAR>
A:Cross-references: GB:I38019; NID:gl464750; PID:g598181
A:Note: authors translated the codon AAG for residue 2325 as Asn
C:Genetics:
A:Gene: GDB:ITPRL
A:Cross-references: GDB:127552; OMIM:147265
A:Map position: 3pter-3p26
C:Superfamily: inositol-trisphosphate receptor

Query Match          60.3%; Score 35; DB 2; Length 2713;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
      | : ||: |||
Db 1319 KAEGKFIKCC 1328

RESULT 11
S54974
type 1 inositol 1,4,5-trisphosphate receptor - human
C:Species: Homo sapiens (man)
C:Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 08-Sep-1997
C:Accession: S54974
R:Yamada, N.; Makino, Y.; Clark, R.A.; Pearson, D.W.; Mattel, M.G.; Guenet, J.L.; Oha
Biochem. J. 302, 781-790, 1994
A:Title: Human inositol 1,4,5-trisphosphate type-1 receptor, InsP(3)R1: structure, fu
A:Reference number: S54974
A:Accession: S54974
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2695 <YAM>
A:Cross-references: EMBL:D26070; NID:g559322; PID:d1005606; PID:g559323
C:Superfamily: inositol-trisphosphate receptor

Query Match          60.3%; Score 35; DB 2; Length 2695;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
      | : ||: |||
Db 1303 KAEGKFIKCC 1312

RESULT 12
B24872
retrovirus-related pol polyprotein homolog - fruit fly (Drosophila melanogaster) retr
C:Species: Drosophila melanogaster
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 08-Jan-1999
C:Accession: B24872
R:Inouye, S.; Yuki, S.; Saigo, K.
Eur. J. Biochem. 154, 417-425, 1986
A:Title: Complete nucleotide sequence and genome organization of a Drosophila transpo
A:Reference number: A91159; MUID:86108354
```

A:Accession: B24872  
A:Molecule type: DNA  
A:Residues: 1-1059 <INO>  
A>Note: the authors translated the codon AAT for residue 1016 as Val and GTA for residue  
C:Genetics:  
A:Gene: FlyBase:297  
A:Cross-references: FlyBase:FBgn0000005  
C:Superfamily: pol polyprotein  
C:Keywords: polyprotein

Query Match 60.3%; Score 35; DB 2; Length 1059;  
Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 KCKFLKK 9  
||:||||  
Db 399 KCEFLKK 405

## RESULT 13

S34639  
pol protein - fruit fly (*Drosophila ananassae*) transposon Tom (fragment)  
C:Species: *Drosophila ananassae*  
C>Date: 22-Nov-1993 #sequence\_revision 01-Dec-1995 #text\_change 24-Sep-1998  
C:Accession: S34639  
R:Tanda, S.; Corces, V.G.  
submitted to the EMBL Data Library, July 1993  
A:Description: Structural and functional analysis of a *Drosophila* retrotransposon that c  
A:Reference number: S34638  
A:Accession: S34639  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1040 <TAN>  
A:Cross-references: EMBL:Z24451; NID:g394703; PID:g394705  
C:Genetics:  
A:Gene: FlyBase:Dana/Tom  
A:Cross-references: FlyBase:FBgn0004357  
A:Mobile element: retrotransposon Tom  
C:Superfamily: pol polyprotein

Query Match 60.3%; Score 35; DB 2; Length 1040;  
Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 KCKFLKK 9  
||:||||  
Db 360 KCEFLKK 366

## RESULT 14

C70028  
conserved hypothetical protein yvan - *Bacillus subtilis*  
C:Species: *Bacillus subtilis*  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998  
C:Accession: C70028  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
A.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.  
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A:Reference number: A69580; MUID:98044033

A:Accession: C70028  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-135 <KUN>  
A:Cross-references: GB:299121; GB:AL009126; NID:g2635827; PID:e1186054; PID:g2635879  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yvan

Query Match 60.3%; Score 35; DB 2; Length 135;  
Best Local Similarity 86.7%; Pred. No. 27;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTKCKFLKK 9  
||:||||  
Db 81 ETKCKLLEK 89

## RESULT 15

S33727  
platelet-derived growth factor receptor alpha precursor - mouse  
C:Species: *Mus musculus* (house mouse)  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Apr-1998  
C:Accession: I57511; S33727  
R:Stiles, C.D.; Wang, C.  
Mol. Cell. Biol. 10, 6781-6784, 1990  
A>Title: Retinoic acid promotes transcription of the platelet-derived growth factor a  
A:Reference number: I57511; MUID:91061789  
A:Accession: I57511  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1089 <RES>  
A:Cross-references: GB:M57683; NID:g199783; PID:g199784  
C:Genetics:  
A:Gene: PDGF-alpha-R  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP; growth factor receptor; transmembrane protein  
F:228-292/Domain: immunoglobulin homology <IMM1>  
F:591-957/Domain: protein kinase homology <KIN>  
F:599-607/Region: protein kinase ATP-binding motif

Query Match 58.6%; Score 34; DB 2; Length 1089;  
Best Local Similarity 71.4%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CKFLKKC 10  
||:||||  
Db 450 CKHIKKC 456

Search completed: September 7, 1999, 23:22:46  
Job time: 953 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:11 ; Search time 71.09 Seconds  
(without alignments)  
3.976 Million cell updates/sec

Title: US-09-124-280A-31

Perfect score: 58

Sequence: 1 KTKCKFLKCC 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	39	67.2	274	1	HAT9_ARATH
2	39	67.2	274	1	HT22_ARATH
3	37	63.8	554	1	NBL4_MOUSE
4	36	62.1	244	1	AEGL1_MOUSE
5	36	62.1	445	1	NEK2_HUMAN
6	35	60.3	2749	1	IP3R_MOUSE
7	35	60.3	2749	1	IP3R_RAT
8	35	60.3	1059	1	POL2_DROME
9	34	58.6	632	1	FBF2_CAEEL
10	34	58.6	403	1	NOF2_DROME
11	34	58.6	984	1	NOF2_DROME
12	34	58.6	553	1	P2B1_YEAST
13	34	58.6	1089	1	PGDS_MOUSE
14	34	58.6	246	1	Y208_METJA
15	34	58.6	206	1	YIM7_YEAST
16	33	56.9	220	1	ANTA_HYDMA
17	33	56.9	318	1	ATH4_ARATH
18	33	56.9	191	1	CC42_DROME
19	33	56.9	282	1	HAT1_ARATH
20	33	56.9	208	1	HAT2_ARATH
21	33	56.9	315	1	HAT3_ARATH
22	33	56.9	284	1	HAT4_ARATH
23	33	56.9	165	1	HT14_ARATH
24	33	56.9	88	1	NAT8_MOUSE
25	33	56.9	3344	1	POLG_PRSVH
26	33	56.9	1002	1	SPHR_AMEPV
27	33	56.9	798	1	YLFN_CAEEL
28	33	56.9	223	1	Y2L2_EBV
29	32	56.0	100	1	TAT_SIVAT
30	32	55.2	1700	1	BAR3_CHITE
31	32	55.2	24	1	BRI_RANBP
32	32	55.2	550	1	CBS_HUMAN
33	32	55.2	560	1	CBS_RAT
34	32	55.2	1035	1	DPOL_RHGM6
35	32	55.2	1097	1	KPCI_CANAL
36	32	55.2	1151	1	KPCI_YEAST
37	32	55.2	226	1	LEF7_NPVAC
38	32	55.2	2499	1	MPRI_BOVIN
39	32	55.2	2491	1	MPRI_HUMAN
40	32	55.2	2483	1	MPRI_MOUSE
41	32	55.2	859	1	MUTS_AQUAE
42	32	55.2	465	1	NRAM_INBLE
43	32	55.2	466	1	NRAM_INBHK

#### ALIGNMENTS

##### RESULT 1

HAT9\_ARATH  
ID HAT9\_ARATH STANDARD; PRT; 274 AA.  
AC P46603;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE HOMEBOX-LEUCINE ZIPPER PROTEIN HAT9 (HD-ZIP PROTEIN 9).  
GN HAT9.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA; VIRIDIPHYTES; EMBRYOPHYTES; TRACHEOPHYTES;  
OC EUPHYLLIPHYTES; SPERMATOPHYTES; MAGNOLIOPHYTES; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPALLOPHITES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV, COLUMBIA;  
RX MEDLINE; 94359937.  
RA SCHEHA M., DAVIS R.W.;  
RT "Structure of homeobox-leucine zipper genes suggests a model for the  
evolution of gene families.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 91:8393-8397(1994).  
CC -|- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
CC -|- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL: U09341; G549892;  
CC EMBL: U09342; G549893;  
CC PROSITE: PS00027; HOMEBOX\_1; 1.  
CC PROSITE: PS50071; HOMEBOX\_2; 1.  
CC PFAM: PF00046; homeobox; 1.  
CC HSP: P01366; IAKH.  
CC HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN.  
FT DNA\_BIND 110 169 HOMEBOX.  
FT DOMAIN 177 198 LEUCINE-ZIPPER.  
FT DOMAIN 90 93 POLY-GLU.  
FT DOMAIN 226 240 POLY-GLY.  
FT CONFLICT 90 90 E -> V (IN G549893).  
SQ SEQUENCE 274 AA; 29867 MW; DA3C3355 CRC32;

Query Match 67.2%; Score 39; DB 1; Length 274;  
Best Local Similarity 85.7%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 CKFLKCC 10

Db 174 CEFLKCC 180

##### RESULT 2

HT22\_ARATH  
ID HT22\_ARATH STANDARD; PRT; 278 AA.  
AC P46604;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE HOMEBOX-LEUCINE ZIPPER PROTEIN HAT22 (HD-ZIP PROTEIN 22).  
GN HAT22.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).

CC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTES; EMBRYOPHYTES; TRACHEOPHYTES;  
 CC EUPHYLLIPHYTES; SPERMATOPHYTES; MAGNOLIOPHYTES; EUDICOTYLEDONS; ROSIDAE;  
 CC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE; 94359937.  
 RA SCHENA M., DAVIS R.W.;  
 RT "Structure of homeobox-leucine zipper genes suggests a model for the  
 evolution of gene families.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:8393-8397(1994).  
 RN [2]  
 RP SEQUENCE OF 166-221 FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE; 92237275.  
 RA SCHENA M., DAVIS R.W.;  
 RT "HD-zip proteins: members of an Arabidopsis homeodomain protein  
 superfamily.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:3894-3898(1992).  
 CC -!- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; U09336; G549887; -;  
 DR EMBL; U09337; G549888; -;  
 DR EMBL; M90417; G166756; -;  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 DR PFAM; PF00046; HOMEBOX; 1.  
 DR HSP; P01366; LYRN.  
 KW HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN.  
 FT DNA\_BIND 123 182 HOMEBOX.  
 FT DOMAIN 190 211 LEUCINE-ZIPPER.  
 FT CONFLICT 221 221 Y -> S (IN REF. 2).  
 SQ SEQUENCE 278 AA; 30729 MW; D50E698D CRC32;  
 Query Match 67.2%; Score 39; DB 1; Length 278;  
 Best Local Similarity 85.7%; Pred. No. 5.2;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 4 CKFLKC 10  
 Db 187 CEFLKC 193  
 RESULT 3  
 ID NBL4\_MOUSE STANDARD; PRT; 554 AA.  
 AC P52963;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE NBL4 PROTEIN.  
 GN EPB4.114.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95074267.  
 RA TAKEUCHI K., KAWASHIMA A., NAGAFUCHI A., TSUKITA S.;  
 RT "Structural diversity of band 4.1 superfamily members.";  
 RL J. CELL SCI. 107:1921-1928(1994).  
 CC -!- TISSUE SPECIFICITY: BRAIN, HEART, LUNG, LIVER AND SPLEEN. NOT  
 DETECTED IN THYMUS AND KIDNEY.  
 CC -----

CC -!- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,  
 RADIAXIN, AND TALIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; D28818; G466548; -;  
 DR MGD; MGI:103007; EPB4.1L4.  
 DR PROSITE; PS00660; BAND\_41\_1; 1.  
 DR PROSITE; PS00661; BAND\_41\_2; 1.  
 DR PROSITE; PS50057; BAND\_41\_3; 1.  
 DR PFAM; PF00373; BAND\_41; 1.  
 KW STRUCTURAL PROTEIN; CYTOSKELETON.  
 FT DOMAIN 70 230 BAND 4.1-LIKE DOMAIN.  
 SQ SEQUENCE 554 AA; 64082 MW; 2057D0C3 CRC32;  
 Query Match 63.8%; Score 37; DB 1; Length 554;  
 Best Local Similarity 70.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 KTKCKFLKCC 10  
 Db 278 KTACKHLWKC 287  
 RESULT 4  
 ID AEG1\_MOUSE STANDARD; PRT; 244 AA.  
 AC Q03401;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE SPERM-COATING GLYCOPROTEIN 1 PRECURSOR (SCP 1) (ACIDIC EPIDIDYMAL  
 GLYCOPROTEIN 1) (CYSTEINE-RICH SECRETORY PROTEIN-1) (CRISP-1).  
 GN AEG1 OR AEG-1.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SUBMANDIBULAR GLAND;  
 RX MEDLINE; 93246016.  
 RA MIZUKI N., KASAHARA M.;  
 RT "Mouse submandibular glands express an androgen-regulated transcript  
 encoding an acidic epididymal glycoprotein-like molecule.";  
 RL MOL. CELL. ENDOCRINOL. 89:25-32(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SUBMANDIBULAR GLAND;  
 RX MEDLINE; 93307144.  
 RA HAENDLER B., KRATZSCHMAR J., THEURING F., SCHLEUNING W.D.;  
 RT "Transcripts for cysteine-rich secretory protein-1 (CRISP-1; DE/AEG)  
 and the novel related CRISP-3 are expressed under androgen control in  
 the mouse salivary gland.";  
 RL ENDOCRINOLOGY 133:192-198(1993).  
 CC -!- FUNCTION: THIS PROTEIN IS SUPPOSED TO HELP SPERMATOZOA UNDERGO  
 FUNCTIONAL MATURATION WHILE THEY MOVE FROM THE TESTIS TO THE  
 DUCTUS DEFERENS.  
 CC -!- SUBCELLULAR LOCATION: STORED IN SECRETORY GRANULES OF GRANULAR  
 CONVOLUTED TUBULES CELLS.  
 CC -!- TISSUE SPECIFICITY: MAINLY FOUND IN THE CAUDA EPIDIDYMIS WHERE IT  
 IS SYNTHESIZED BY THE PRINCIPAL CELLS AND SECRETED INTO THE LUMEN.  
 CC BINDS TO THE HEADS OF SPERMATOZOA. ALSO EXPRESSED IN THE  
 SUBMANDIBULAR GLAND.  
 CC -!- DEVELOPMENTAL STAGE: EXPONENTIAL INCREASE BETWEEN DAYS 25 AND 30  
 AFTER BIRTH.  
 CC -!- INDUCTION: THIS PROTEIN IS ANDROGEN-DEPENDENT.

CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M92849; GI91738; -  
 CC EMBL: L05559; G309191; -  
 CC PIR: A49202; A49202.  
 CC DR MGD; MGI:102553; AEG1.  
 CC DR PROSITE; PS01009; SCP\_AG5\_PRI\_SC7\_1; 1.  
 CC DR PROSITE; PS01010; SCP\_AG5\_PRI\_SC7\_2; 1.  
 CC DR PFAM; PF00188; SCP; 1  
 CC KW SPERM; GLYCOPROTEIN; SIGNAL.  
 CC FT SIGNAL  
 CC FT CHAIN 1  
 CC FT CHAIN 20 244 SPERM-COATING GLYCOPROTEIN 1.  
 CC FT CARBOHYD 145 145 POTENTIAL.  
 CC SQ SEQUENCE 244 AA; 27679 MW; BFBD86F2 CRC32;  
 CC  
 CC Query Match 62.1%; Score 36; DB 1; Length 244;  
 CC Best Local Similarity 75.0%; Pred. No. 15;  
 CC Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TKCKFLKK 9

| | | | |

Db 213 TNCKYLKK 220

RESULT 5

ID NEK2\_HUMAN STANDARD; PRT; 445 AA.  
 AC P51955;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE SERINE/THREONINE-PROTEIN KINASE NEK2 (EC 2.7.1.1-) (NIMA-RELATED  
 DE PROTEIN KINASE 2) (NIMA-LIKE PROTEIN KINASE 1) (HSPK 21).  
 GN NEK2 OR NLK1.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-T-CELL, PLACENTA, AND NASOPHARYNX;  
 RX MEDLINE; 94368699.  
 RA SCHULTZ S.J., FRY A.M., SUETTERLIN C., RIED T., NIGG E.A.;  
 RT "Cell cycle-dependent expression of Nek2, a novel human protein  
 RT kinase related to the NIMA mitotic regulator of Aspergillus  
 RT nidulans.";  
 RL CELL GROWTH DIFFER. 5:625-635(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA LU K.P., HUNTER T.;  
 RL SUBMITTED (JUL-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE OF 83-203 FROM N.A.  
 RX MEDLINE; 94100173.  
 RA SCHULTZ S.J., NIGG E.A.;  
 RT "Identification of 21 novel human protein kinases, including 3  
 RT members of a family related to the cell cycle regulator nima of  
 RT Aspergillus nidulans.";  
 RL CELL GROWTH DIFFER. 4:821-830(1993).  
 CC -1- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION.  
 CC MAY HAVE A ROLE AT THE G2-M TRANSITION.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES THROUGHOUT S PHASE AND SHOWS  
 CC MAXIMAL LEVELS IN LATE G2. THIS EXPRESSION PATTERN IS HIGHLY

CC REMINISCENT OF THAT OF A AND B CYCLINS.  
 CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-  
 CC PROTEIN KINASES. BELONGS TO THE NIMA SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z29066; G479171; -  
 CC EMBL: U11050; G507875; -  
 CC DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 CC DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC DR PFAM; PF00069; pkinase; 1.  
 CC DR HSP; P24941; LAQ1.  
 CC KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING; MITOSIS;  
 CC NUCLEAR PROTEIN; PHOSPHORYLATION; CELL CYCLE; CELL DIVISION.  
 CC FT DOMAIN 8 271 PROTEIN KINASE.  
 CC FT NP\_BIND 14 22 ATP (BY SIMILARITY).  
 CC FT BINDING 37 37 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 141 141 BY SIMILARITY.  
 CC FT MOD\_RES 179 179 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 CC FT CONFLICT 84 85 IV -> LV (IN REF. 2).  
 CC SQ SEQUENCE 445 AA; 51763 MW; 86198471 CRC32;

Query Match 62.1%; Score 36; DB 1; Length 445;

Best Local Similarity 77.8%; Pred. No. 25;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTKCKFLKK 9

| | | | |

Db 407 KSKCKDLKK 415

RESULT 6

ID IP3R\_MOUSE STANDARD; PRT; 2749 AA.  
 AC P11881;  
 DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN TYPE 1 RECEPTOR  
 DE (INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN P(400)) (TYPE 1 INSP3  
 DE RECEPTOR).  
 GN INSP3R OR ITPRI.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PURKINJE CELLS;  
 RX MEDLINE; 90044039.  
 RA FURUICHI T., YOSHIKAWA S., MIYAWAKI A., WADA K., MAEDA N.,  
 RA MIKOSHIBA K.;  
 RT "Primary structure and functional expression of the inositol 1,4,5-  
 RT trisphosphate-binding protein p400.";  
 RL NATURE 342:32-38(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ICR; TISSUE=CEREBELLUM;  
 RX MEDLINE; 89345101.  
 RA FURUICHI T., YOSHIKAWA S., MIKOSHIBA K.;  
 RT "Nucleotide sequence of cDNA encoding p400 protein in the mouse  
 RT cerebellum.";  
 RL NUCLEIC ACIDS RES. 17:5385-5386(1989).  
 RN [3]  
 RP SEQUENCE OF 318-332 AND 1692-1731 FROM N.A., AND ALTERNATIVE SPLICING.

RC STRAIN-JCR;  
RX MEDLINE; 91296797.  
RA NAKAGAWA T., OKANO H., FURUICHI T., ARUGA J., MIKOSHIBA K.;  
RT "The subtypes of the mouse inositol 1,4,5-trisphosphate receptor are  
RT expressed in a tissue-specific and developmentally specific manner.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 88:6244-6248(1991).  
CC -!- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND  
CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
CC THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL  
CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-  
CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE  
CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
CC RETICULUM.  
CC -!- PTM: PHOSPHORYLATED BY CYCLIC-AMP KINASE. PHOSPHORYLATION PREVENTS  
CC THE LIGAND-INDUCED OPENING OF THE CALCIUM CHANNELS.  
CC -!- CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE RECEPTOR, MOST  
CC PROBABLY BY INTERACTING WITH A DISTINCT CALCIUM-BINDING PROTEIN  
CC WHICH THEN INHIBITS THE RECEPTOR.  
CC -!- ALTERNATIVE PRODUCTS: ADDITIONAL SUBTYPES OF INSP3R ARISE BY  
CC ALTERNATIVE SPLICING OF THE SAME GENE.  
CC -!- SIMILARITY: TO RYANODINE RECEPTOR.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X15373; G53569; .  
CC DR EMBL; M75986; G198412; .  
CC DR EMBL; M75987; G554171; .  
CC DR PIR; S04844; ACMSIT.  
CC DR MGI; 96623; ITPRI.  
CC KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; PHOSPHORYLATION;  
CC ENDOPLASMIC RETICULUM; IONIC CHANNEL; ION TRANSPORT; CALCIUM CHANNEL;  
CC ALTERNATIVE SPLICING.  
CC FT DOMAIN 1 2273  
CC FT TRANSMEM 2274 2294  
CC FT TRANSMEM 2308 2326  
CC FT TRANSMEM 2334 2356  
CC FT TRANSMEM 2365 2387  
CC FT TRANSMEM 2391 2407  
CC FT TRANSMEM 2440 2462  
CC FT TRANSMEM 2530 2549  
CC FT TRANSMEM 2570 2589  
CC FT DOMAIN 2590 2749  
CC FT VARSPLIC 318 332  
CC FT VARSPLIC 1692 1731  
CC FT VARSPLIC 1715 1715  
CC FT VARSPLIC 1715 1731  
CC FT MOD\_RES 1588 1588  
CC FT MOD\_RES 1755 1755  
CC SQ SEQUENCE 2749 AA; 313193 MW; B19B78F2 CRC32;  
  
Query Match 60.38; Score 35; DB 1; Length 2749;  
Best Local Similarity 60.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 KTKCKFLKCC 10  
Db 1317 KAEGKFIKK 1326  
  
RESULT 7  
IP3R\_RAT  
ID IP3R\_RAT STANDARD; PRT; 2749 AA.  
AC P29994;  
DT 01-APR-1993 (REL. 25, CREATED)

DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
DE 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN TYPE 1 RECEPTOR  
DE (INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN P(400)) (TYPE 1 INSP3  
DE RECEPTOR).  
GN INSP3R.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE; 90324264.  
RA MIGNERY G.A., NEWTON C.L., ARCHER B.T. III, SUEHOF T.C.;  
RT "Structure and expression of the rat inositol 1,4,5-trisphosphate  
RT receptor.";  
RL J. BIOL. CHEM. 265:12679-12685(1990).  
[2]  
RN SEQUENCE OF 1652-1824 FROM N.A., AND ALTERNATIVE SPLICING.  
RC TISSUE-BRAIN;  
RX MEDLINE; 91187909.  
RA DANOFF S.K., FERRIS C.D., DONATH C., FISCHER G.A., MUNEMITSU S.,  
RA ULLRICH A., SNYDER S.H., ROSS C.A.;  
RT "Inositol 1,4,5-trisphosphate receptors: distinct neuronal and  
RT nonneuronal forms derived by alternative splicing differ in  
RT phosphorylation.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 88:2951-2955(1991).  
CC -!- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND  
CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
CC THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL  
CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-  
CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE  
CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
CC RETICULUM.  
CC -!- TISSUE SPECIFICITY: THE LONGER ISOFORM IS FOUND IN THE BRAIN WHILE  
CC THE SHORTER ISOFORM IS FOUND IN THE FETAL BRAIN AND PERIPHERAL  
CC TISSUES.  
CC -!- PTM: PHOSPHORYLATED BY CYCLIC-AMP KINASE. PHOSPHORYLATION PREVENTS  
CC THE LIGAND-INDUCED OPENING OF THE CALCIUM CHANNELS.  
CC -!- CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE RECEPTOR, MOST  
CC PROBABLY BY INTERACTING WITH A DISTINCT CALCIUM-BINDING PROTEIN  
CC WHICH THEN INHIBITS THE RECEPTOR.  
CC -!- ALTERNATIVE PRODUCTS: TWO FORMS, A LONG (NEURONAL) AND A SHORT  
CC (NONNEURONAL) FORM, ARE PRODUCED BY ALTERNATIVE SPLICING OF THE  
CC SAME GENE.  
CC -!- SIMILARITY: TO RYANODINE RECEPTOR.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J05510; G204675; .  
CC DR EMBL; J05510; G204674; .  
CC DR EMBL; M64699; G204967; .  
CC DR EMBL; M64698; G204969; .  
CC DR PIR; A36579; A36579.  
CC KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; PHOSPHORYLATION;  
CC ENDOPLASMIC RETICULUM; IONIC CHANNEL; ION TRANSPORT; CALCIUM CHANNEL;  
CC ALTERNATIVE SPLICING.  
CC FT DOMAIN 1 2273  
CC FT TRANSMEM 2274 2294  
CC FT TRANSMEM 2308 2326  
CC FT TRANSMEM 2334 2356  
CC FT TRANSMEM 2365 2387  
CC FT TRANSMEM 2391 2407  
CC FT TRANSMEM 2440 2462  
CC FT TRANSMEM 2530 2549  
CC FT TRANSMEM 2570 2589



```

FT DOMAIN 2590 2749 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 1589 1589 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT MOD_RES 1755 1755 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT VARIANT 1372 1372 MISSING (IN ALL, BUT P17 CLONES).
FT VARSPLIC 322 336 MISSING (IN VARIANT).
FT VARSPLIC 1693 1731 MISSING (IN SHORT FORM).
FT CONFLICT 1716 1716 E -> EQ (IN REF. 2).
FT CONFLICT 1763 1763 P -> R (IN REF. 2).
SQ SEQUENCE 2749 AA; 313132 MW; A64DE4B1 CRC32;

Query Match 60.38; Score 35; DB 1; Length 2749;
Best Local Similarity 60.08; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
   | : |||
Db 1317 KAEGKFIKCC 1326

RESULT 8
POL2_DROME STANDARD; PRT; 1059 AA.
AC P20825; 1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE RETROVIRUS-RELATED POL PROTEIN FROM TRANSPOSON 297 [CONTAINS:
DE PROTEASE (EC 3.4.23.-); REVERSE TRANSCRIPTASE (EC 2.7.7.49);
DE ENDONUCLEASE).
GN POL.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86108354.
RA INOUE S., YUKI S., SAIGO K.;
RT "Complete nucleotide sequence and genome organization of a Drosophila
transposable genetic element, 297."
RL EUR. J. BIOCHEM. 154:417-425(1986).
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY U22.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X03431; -; NOT_ANNOTATED_CDS.
CC PIR; B24872; B24872.
CC DROSOPHILA MELANOGASTER (FRUIT FLY).
CC FLYBASE; FBgn0000005; 297.
CC PROSITE; PS00141; ASP_PROTEASE; 1.
CC PFAM; PF00077; rvp; 1.
CC PFAM; PF00078; rvt; 1.
CC HSP; P04585; lrti.
CC HYDROLASE; ASPATYL PROTEASE; RNA-DIRECTED DNA POLYMERASE;
KW ENDONUCLEASE; TRANSFERASE; POLYPROTEIN; TRANSPOSABLE ELEMENT.
FT ACT_SITE 30 30 PROTEASE (BY SIMILARITY).
SQ SEQUENCE 1059 AA; 123310 MW; 78A65BF9 CRC32;

Query Match 60.38; Score 35; DB 1; Length 1059;
Best Local Similarity 85.78; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 KCKFLKK 9
   |||
Db 399 KCEFLKK 405

us-09-124-280a-31.rsp

```

```

RESULT 9
FBF2_CAEEL STANDARD; PRT; 632 AA.
AC Q09312;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE FBF-2 PROTEIN.
GN FBF-2 OR F2H12.5.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA FAVELLO T.;
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: C.ELEGANS F54C9.8, YEAST YGL014W, S.POMBE SPAC4G8.03C
AND D.MELANOGASTER PUMILLO.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U23176; G726408; -;
CC DROSOPHILA MELANOGASTER (FRUIT FLY).
CC DROSOPHILA MELANOGASTER (FRUIT FLY).
CC PTERYGOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
CC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
CC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OREGON-R;
RX MEDLINE; 89356666.
RA TEMPLETON N.S., POTTER S.S.;
RT "Complete foldback transposable elements encode a novel protein found
in Drosophila melanogaster."
RL EMBO J. 8:1887-1894(1989).
CC -!- FUNCTION: NOT KNOWN.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- SIMILARITY: TO C-TERMINAL OF NOF-FB ELEMENT PROTEIN IN
CC STRAIN TE146(Z).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

Query Match 58.68; Score 34; DB 1; Length 632;
Best Local Similarity 56.78; Pred. No. 72;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTKCKFLKK 9
   |||
Db 201 KTGCFLEK 209

RESULT 10
NOF2_DROME STANDARD; PRT; 403 AA.
AC P15297;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE 46 KD PROTEIN IN NOF-FB TRANSPOSABLE ELEMENT.
GN NOF.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OREGON-R;
RX MEDLINE; 89356666.
RA TEMPLETON N.S., POTTER S.S.;
RT "Complete foldback transposable elements encode a novel protein found
in Drosophila melanogaster."
RL EMBO J. 8:1887-1894(1989).
CC -!- FUNCTION: NOT KNOWN.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- SIMILARITY: TO C-TERMINAL OF NOF-FB ELEMENT PROTEIN IN
CC STRAIN TE146(Z).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```



```
CC EMBL; M64839; G171149; -
DR EMBL; X66490; G3547; -
DR EMBL; X54963; G3558; -
DR EMBL; U21094; G665973; -
DR PIR; S16809; S16809
DR PIR; S21186; S21186
DR PIR; A41152; A41152
DR SGD; L0000370; CNA1
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; 1
DR PFAM; PF00149; Stposphatase; 1
DR HSP; Q08209; IAU1
DR HYDROLASE; CALMODULIN-BINDING; IRON; MANGANESE; MULTIGENE FAMILY.
KW VARIANT 45 46 PI -> SY (IN REF. 1)
SQ SEQUENCE 553 AA; 63002 MW; DF67C62F CRC32;

Query Match 58.68; Score 34; DB 1; Length 553;
Best Local Similarity 75.08; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLK 8
Db 299 KASCKFLK 306
1 |||||

RESULT 13
PGDS_MOUSE STANDARD; PRT; 1089 AA.
AC P26618; Q62046;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE ALPHA PLATELET-DERIVED GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112)
DE (PDGF-R-ALPHA).
GN PDGFRA.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91061789.
RA STILES C.D., WANG C.;
RT "Retinoic acid promotes transcription of the platelet-derived growth
RT factor alpha-receptor gene."
RL MOL. CELL. BIOL. 10:6781-6784(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92334866.
RA DO M.S., FITZER-ATTAS C., GUBBAY J., GREENFELD L., FELDMAN M.,
RA EISENBACH L.;
RT "Mouse platelet-derived growth factor alpha receptor: sequence,
RT tissue-specific expression and correlation with metastatic
RT phenotype."
RL ONCOGENE 7:1567-1575(1992).
CC -1- FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND
CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR CAN BIND
CC EITHER PDGF-A OR PDGF-B.
CC -1- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA
CC SUBUNITS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. THE
CC EXTRACELLULAR DOMAIN CONTAINS FIVE IG-FOLD DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; M84607; G200275; -
DR EMBL; M57683; G199784; -
DR PIR; S33727; S33727.
DR MGD; MGI:97530; PDGFRA.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00047; Ig; 2.
DR PFAM; PF00069; pkinase; 2.
DR HSP; P11362; IAGW.
DR TYROSINE-PROTEIN KINASE; RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW TRANSFERASE; PHOSPHORYLATION; ATP-BINDING; IMMUNOGLOBULIN FOLD;
KW SIGNAL.
FT CHAIN 1 24 POTENTIAL.
FT SIGNAL 25 1089 ALPHA PLATELET-DERIVED GROWTH FACTOR
FT DOMAIN 25 525 RECEPTOR.
FT TRANSMEM 526 549 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 550 1089 POTENTIAL.
FT DOMAIN 593 954 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 599 607 PROTEIN KINASE.
FT BINDING 627 627 ATP (BY SIMILARITY).
FT ACT_SITE 818 818 ATP (BY SIMILARITY).
FT CARBOHYD 42 42 BY SIMILARITY.
FT CARBOHYD 76 76 POTENTIAL.
FT CARBOHYD 89 89 POTENTIAL.
FT CARBOHYD 103 103 POTENTIAL.
FT CARBOHYD 179 179 POTENTIAL.
FT CARBOHYD 353 353 POTENTIAL.
FT CARBOHYD 359 359 POTENTIAL.
FT CARBOHYD 458 468 POTENTIAL.
FT CARBOHYD 468 468 POTENTIAL.
FT CARBOHYD 506 506 POTENTIAL.
FT CONFLICT 192 192 T -> A (IN REF. 1).
FT CONFLICT 202 202 E -> A (IN REF. 1).
FT CONFLICT 252 252 E -> G (IN REF. 1).
FT CONFLICT 271 271 L -> V (IN REF. 1).
FT CONFLICT 322 322 G -> S (IN REF. 1).
FT CONFLICT 326 326 A -> P (IN REF. 1).
FT CONFLICT 439 440 GT -> EG (IN REF. 1).
FT CONFLICT 529 529 A -> E (IN REF. 1).
FT CONFLICT 737 737 A -> D (IN REF. 1).
FT CONFLICT 849 849 Y -> D (IN REF. 1).
FT CONFLICT 936 936 E -> D (IN REF. 1).
FT CONFLICT 950 950 V -> L (IN REF. 1).
FT CONFLICT 1005 1005 S -> G (IN REF. 1).
SQ SEQUENCE 1089 AA; 122727 MW; 82FAE940 CRC32;

Query Match 58.68; Score 34; DB 1; Length 1089;
Best Local Similarity 71.4%; Pred. No. 11e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CKFLKCC 10
Db 450 CKHKCC 456
|| :|||

RESULT 14
Y208_METJA STANDARD; PRT; 246 AA.
AC Q57661;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0208.
GN MJ0208.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE: 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii".
RL SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC -!- SIMILARITY: HIGH, TO M.JANNASCHII MJ0730.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67476; G1498983; -
DR TIGR: MJ0208; -
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 2.
DR PFAM: PF00037; fer4; 1.
DR HSP: P00208; 1BLU.
KW HYPOTHETICAL PROTEIN; ELECTRON TRANSPORT; IRON-SULFUR; 4FE-4S.
SQ SEQUENCE 246 AA; 27696 MW; 569E220E CRC32;

Query Match
Best Local Similarity 58.6%; Score 34; DB 1; Length 246;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 10
DB 156 KKKCKLCLKC 165

RESULT 15
YIM7_YEAST
ID YIM7_YEAST STANDARD; PRT; 206 AA.
AC P40470;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 23.8 KD PROTEIN IN MET18-STH INTERGENIC REGION.
GN Y11127C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA BARRELL B.G., BADCOCK K., BANKIER A.T., BOWMAN S., BROWN D.,
RA CHURCHER C.M., CONNOR R., COPSEY T., DEAR S., DEVLIN K., FRASER A.,
RA GENTLES S., HAMLYN N., HORSNELL T.S., HUNT S., JAGELS K., JONES M.,
RA LOUIS E., LYE G., MOULE S., MOULE T., ODELL C., PEARSON D.,
RA RAJANDREAM M.A., RILES L., ROWLEY N., SKELTON J., SMITH V.,
RA WALSH S.V., WHITEHEAD S.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 247047; G763219; -
DR EMBL: 246833; G599990; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 206 AA; 23845 MW; A140E520 CRC32;

Query Match
Best Local Similarity 58.6%; Score 34; DB 1; Length 206;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKCC 9
DB 68 KKKCKLVKK 76

Search completed: September 7, 1999, 23:59:12
Job time: 524 sec
```

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:41 ; Search time 160.2 Seconds  
(without alignments)  
3.842 Million cell updates/sec

Title: US-09-124-280A-31

Perfect score: 58

Sequence: 1 KTKCKFLKKC 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL10:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	39	67.2	274	10	O82763 arabidopsis
2	38.5	66.4	239	12	O55741 chilo iride
3	37	63.8	971	5	O45785 caenorhabdi
4	37	63.8	125	12	O68540 horseradish
5	36	62.1	347	2	O51533 borrelia bu
6	36	62.1	583	5	O27472 caenorhabdi
7	36	62.1	284	10	O39862 glycine max
8	36	62.1	319	10	O40780 pimpinella
9	36	62.1	318	10	O40781 pimpinella
10	36	62.1	302	10	O40782 pimpinella
11	36	62.1	171	10	O43441 glycine max
12	36	62.1	422	12	O9YVW4 melanoplus
13	36	62.1	1009	12	O9YVTO4 ateline her
14	35	60.3	135	2	O32236 bacillus su
15	35	60.3	2713	4	O14660 homo sapien
16	35	60.3	2695	4	O14643 homo sapien
17	35	60.3	151	5	O15655 plasmodium
18	35	60.3	294	5	O18972 caenorhabdi
19	35	60.3	125	5	P91462 caenorhabdi
20	35	60.3	486	5	O01822 caenorhabdi
21	35	60.3	1139	5	O97242 plasmodium
22	34	58.6	298	1	O27241 methanobact
23	34	58.6	189	2	O67915 aquifex aeo
24	34	58.6	60	2	O92383 escherichia
25	34	58.6	1121	3	O12734 saccharomyc
26	34	58.6	603	4	O75135 homo sapien
27	34	58.6	548	5	O19496 caenorhabdi
28	34	58.6	269	5	O18632 caenorhabdi
29	34	58.6	338	5	O17928 caenorhabdi

30	34	58.6	200	5	O23170 caenorhabdi
31	34	58.6	504	11	Q60565 mesocricatu
32	34	58.6	464	11	Q921K5 mus musculu
33	33.5	57.8	1070	3	P78734 pneumocysti
34	33	56.9	163	1	P81292 methanococ
35	33	56.9	170	1	O59556 pyrococcus
36	33	56.9	355	1	O58566 methanococ
37	33	56.9	492	2	O25029 helicobacte
38	33	56.9	382	2	O85003 streptococ
39	33	56.9	388	2	O54685 lactococcc
40	33	56.9	383	2	O48621 lactococcc
41	33	56.9	491	2	O92MI8 helicobacte
42	33	56.9	217	2	Q92AC5 campylobact
43	33	56.9	144	2	Q929K7 bacillus sp
44	33	56.9	301	3	Q06701 saccharomyc
45	33	56.9	1447	11	P70365 mus musculu

#### ALIGNMENTS

```

RESULT 1
O82763
ID O82763 PRELIMINARY: PRT: 274 AA.
AC O82763;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE HOMEBOX PROTEIN, HAT9.
GN T30L20.6 OR T20K9.1.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,
RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;
RT "Arabidopsis thaliana chromosome II BAC T30L20 genomic sequence.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,
RA SYKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,
RA SOMERVILLE C.R., VENTER J.C.;
RL "Arabidopsis thaliana chromosome II BAC T20K9 genomic sequence.";
RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AC005617; AAC63591.1; -;
DR EMBL; AC004786; AAC32427.1; -;
DR PFAM; PF00046; homeobox; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
KW Homeobox; DNA-binding; Nuclear protein.
SQ SEQUENCE 274 AA; 29877 MW; 0797F7C1 CRC32;

```

Query Match 67.2%; Score 39; DB 10; Length 274;  
Best Local Similarity 85.7%; Pred. No. 7.9;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 CKFLKKC 10  
|:|||||

Db 174 CEFLKKC 180

RESULT 2

O55741 PRELIMINARY: PRT: 239 AA.  
ID O55741;  
AC O55741;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last annotation update)  
 DE HYPOTHETICAL 28.0 KD PROTEIN.  
 OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).  
 OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA BAHU U., TIDONA C.A., DARAI G.;  
 RL Virus Genes 0:0-0(1997).  
 DR EMBL: AF003534; AAB94452.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 239 AA; 27379 MW; 7B15C915 CRC32;

Query Match 66.4%; Score 38.5; DB 12; Length 239;  
 Best Local Similarity 63.6%; Pred. No. 8.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

Oy 1 KTK-CKFLKCC 10  
 ||: ||: |||  
 Db 83 KTRMCKYVKKC 93

## RESULT 3

ID O45785 PRELIMINARY; PRT; 971 AA.  
 AC O45785; O45816;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-JAN-1999 (TReMBLrel. 09, Last annotation update)  
 DE T12D8.9 PROTEIN.  
 GN T12D8.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditia; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA GARDNER A.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAN S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKIN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MCMURRAY A.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; 283241; CAB05819.1; -.  
 DR EMBL; 281120; CAB05819.1; JOINED.  
 DR EMBL; 281120; CAB03350.1; -.  
 DR EMBL; 283241; CAB03350.1; JOINED.  
 SQ SEQUENCE 971 AA; 113043 MW; 5FF79B6D CRC32;

Query Match 63.8%; Score 37; DB 5; Length 971;  
 Best Local Similarity 60.0%; Pred. No. 46;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KTKCKFLKCC 10  
 ||: ||: |||  
 Db 341 KTLCKYRKSC 350

## RESULT 4

ID Q68540 PRELIMINARY; PRT; 125 AA.  
 AC Q68540;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE MOVEMENT PROTEIN V2.  
 GN V2.  
 OS Horseradish curly top virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96335668.  
 RA KLUTE K.A., NADLER S.A., STENGER D.C.;  
 RT "Horseradish curly top virus is a distinct subgroup II geminivirus  
 species with rep and C4 genes derived from a subgroup III ancestor."  
 RL J. Gen. Virol. 77:1369-1378(1996).  
 DR EMBL; U49907; AAB18923.1; -.  
 SQ SEQUENCE 125 AA; 14506 MW; 580ED7E5 CRC32;

Query Match 63.8%; Score 37; DB 12; Length 125;  
 Best Local Similarity 60.0%; Pred. No. 9.9;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KTKCKFLKCC 10  
 ||: ||: |||  
 Db 62 KRKCTFSKCC 71

## RESULT 5

ID O51533 PRELIMINARY; PRT; 347 AA.  
 AC O51533;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE FEMA PROTEIN (FEMA).  
 GN BR0586.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUOT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 burgdorferi";  
 RL Nature 390:580-586(1997).  
 DR EMBL; AE001160; AAB91521.1; -.  
 DR TIGR; BR0586; -.  
 SQ SEQUENCE 347 AA; 41521 MW; BA5C09E8 CRC32;

Query Match 62.1%; Score 36; DB 2; Length 347;  
 Best Local Similarity 77.8%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KTKCKFLKCC 9  
 ||: ||: |||  
 Db 121 KTKIKYLKCC 129

RESULT 6  
Q27472 PRELIMINARY; PRT; 583 AA.  
AC Q27472;  
DT 01-NOV-1996 (TEMBLrel. 01, Created)  
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
DE SIMILAR TO CYTOCHROME P450. NCBI GI: 1255383 (EC 1.14.14.1).  
GN C26F1.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA GRISEL C., STELLYES L., BRADSHAW H.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA WATERSTON R.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA WATERSTON R.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; U53148; AAB37074.1; --  
DR PFAM; PF00067; P450.1;  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.  
FT BINDING 485 HEME (BY SIMILARITY).  
SQ SEQUENCE 583 AA; 67366 MW; AEB7DD4C CRC32;

Query Match 62.1%; Score 36; DB 5; Length 583;  
Best Local Similarity 62.5%; Pred. No. 47;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KCKFLKCC 10  
||:|:|  
DB 390 KCSYLEKC 397

RESULT 7  
Q39862 PRELIMINARY; PRT; 284 AA.  
AC Q39862;  
DT 01-NOV-1996 (TEMBLrel. 01, Created)  
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)  
DE HOMEBOX-LEUCINE ZIPPER PROTEIN.  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;

OC Glycine.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. HWANG-KEUM;  
RA MOON Y.H., CHOI S.B., KIM J.I., HAN T.J., CHO S.H., KIM W.T.,  
RA LEE K.W.;  
RL Mol. Cells 6:366-373(1996).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL; X92489; CAA63222.1; --  
DR MENDEL; 8291; GLYma; I345.1;  
DR PFAM; PF00046; homeobox; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
KW Homeobox; DNA-binding; Nuclear protein; Homeobox.  
SQ SEQUENCE 284 AA; 31919 MW; 9D3F8C00 CRC32;

Query Match 62.1%; Score 36; DB 10; Length 284;  
Best Local Similarity 71.4%; Pred. No. 27;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 CKFLKCC 10  
|:|:|:|  
DB 175 CEFLKRC 181

RESULT 8  
Q40780 PRELIMINARY; PRT; 319 AA.  
AC Q40780;  
DT 01-NOV-1996 (TEMBLrel. 01, Created)  
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)  
DE HOMEBOX-LEUCINE ZIPPER PROTEIN.  
OS Pimpinella brachycarpa.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Asteridae; euasterids II; Apiales; Apiaceae;  
OC Pimpinella.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-SHOOT APEX;  
RA MOON Y.H., CHOI S.B., KIM J.I., KIM J.C., HAN T.J., CHO S.H.,  
RA LEE K.W.;  
RL Mol. Cells 6:697-703(1996).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL; X94375; CAA64152.1; --  
DR MENDEL; I2126; PIMbr; I345.1;  
DR PFAM; PF00046; homeobox; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
KW Homeobox; DNA-binding; Nuclear protein; Homeobox.  
SQ SEQUENCE 319 AA; 36055 MW; D71BE66A CRC32;

Query Match 62.1%; Score 36; DB 10; Length 319;  
Best Local Similarity 71.4%; Pred. No. 30;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 CKFLKCC 10  
|:|:|:|  
DB 214 CEFLKRC 220

RESULT 9  
Q40781 PRELIMINARY; PRT; 318 AA.  
AC Q40781;  
DT 01-NOV-1996 (TEMBLrel. 01, Created)  
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)  
DE HOMEBOX-LEUCINE ZIPPER PROTEIN.  
OS Pimpinella brachycarpa.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;

OC core eudicots; Asteridae; euasterids II; Apiales; Apiaceae;  
 CC Pimpinella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SHOOT APEX;  
 RA MOON Y.H., CHOI S.B., KIM J.I., KIM J.C., HAN T.J., CHO S.H.,  
 RA LEE K.W.;  
 RL MOL. Cells 6:697-703(1996).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC EMBL: X95193; CAA64491.1; -;  
 DR MENDEL: 12127; PIMbr:1345;2.  
 DR PFAM: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Homeobox.  
 SQ SEQUENCE 318 AA; 35936 MW; CB5B0D6F CRC32;

Query Match 62.1%; Score 36; DB 10; Length 318;  
 Best Local Similarity 71.4%; Pred. No. 30;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CKFLKCC 10  
 I:||||  
 Db 213 CEFLKRC 219

## RESULT 10

ID Q40782 PRELIMINARY; PRT; 302 AA.  
 AC Q40782;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE HOMEBOX-LEUCINE ZIPPER PROTEIN.  
 GN PH24.  
 OS Pimpinella brachycarpa.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 CC core eudicots; Asteridae; euasterids II; Apiales; Apiaceae;  
 RN Pimpinella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SHOOT APEX;  
 RA MOON Y.H., CHOI S.B., KIM J.I., KIM J.C., HAN T.J., CHO S.H.,  
 RA LEE K.W.;  
 RL MOL. Cells 6:697-703(1996).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: X94449; CAA64221.1; -;  
 DR MENDEL: 12128; PIMbr:1345;3.  
 DR PFAM: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Homeobox.  
 SQ SEQUENCE 302 AA; 33664 MW; A8699724 CRC32;

Query Match 62.1%; Score 36; DB 10; Length 302;  
 Best Local Similarity 71.4%; Pred. No. 29;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CKFLKCC 10  
 I:||||  
 Db 200 CEFLKRC 206

## RESULT 11

ID Q43441 PRELIMINARY; PRT; 171 AA.  
 AC Q43441;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE HOMEBOX-LEUCINE ZIPPER PROTEIN (FRAGMENT).  
 OS Glycine max (soybean).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 CC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 RN Glycine.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. WHANG-KHEUM; TISSUE=GLOBAL SOMATIC EMBRYO;  
 RA MOON Y.H., CHOI S.B., LEE K.W.;  
 RL SUBMITTED (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: U30475; AAA74017.1; -;  
 DR MENDEL: 8292; GLYma:1345;2.  
 DR PFAM: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 171 AA; 15903 MW; 261D385C CRC32;

Query Match 62.1%; Score 36; DB 10; Length 171;  
 Best Local Similarity 71.4%; Pred. No. 19;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CKFLKCC 10  
 I:||||  
 Db 62 CEFLKRC 68

## RESULT 12

ID Q9YVW4 PRELIMINARY; PRT; 422 AA.  
 AC Q9YVW4;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE ORF MSV128 HYPOTHETICAL PROTEIN.  
 GN MSV128.  
 OS Melanoplus sanguinipes entomopoxvirus.  
 CC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;  
 CC Entomopoxvirus B.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TUCSON;  
 RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.;  
 RT "The genome of Melanoplus sanguinipes entomopoxvirus.";  
 RL J. Virol. 73:533-552(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TUCSON;  
 RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF063866; AAC97795.1; -;  
 SQ SEQUENCE 422 AA; 50633 MW; 93A75459 CRC32;

Query Match 62.1%; Score 36; DB 12; Length 422;  
 Best Local Similarity 85.7%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KCFLKK 9  
 I:||||  
 Db 293 KCFLKK 299

## RESULT 13

ID Q9YTQ4 PRELIMINARY; PRT; 1009 AA.  
 AC Q9YTQ4;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE DNA POLYMERASE (EC 2.7.7).  
 OS Ateline herpesvirus 3.



OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae: Rhadinovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=73;  
 RA ALBRECHT J.-C., FLECKENSTEIN B.;  
 RT "Primary Structure of the Herpesvirus Ateles Genome.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N  
 CC PYROPHOSPHATE + DNA(N).  
 DR EMBL: AF083424; AAC95533.1; -;  
 DR PROSITE: PS00116; DNA-POLYMERASE B; 1.  
 KW DNA-directed DNA polymerase; DNA replication; DNA-binding.  
 SQ SEQUENCE 1009 AA; 114356 MW; 8F89203 CRC32;

Query Match 62.1%; Score 36; DB 12; Length 1009;  
 Best Local Similarity 66.7%; Pred. No. 70;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTKCKFLKK 9  
 ||| |||:  
 DB 815 KTKCKFLKK 823

RESULT 14  
 O32236

ID O32236 PRELIMINARY; PRT; 135 AA.  
 AC O32236;  
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)  
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)  
 DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)  
 DE YVAN PROTEIN.  
 GN YVAN.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE: 98044033.  
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AVEDEO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNEL S.C., BRON S.,  
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRICH S.D., EMERSON P.T.,  
 RA ENTIAN K.D., ERLINGTON J., FABRET C., FERRARI E., FOULGER D.,  
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,  
 RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
 RA GUISEPI G., GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A.,  
 RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,  
 RA JORIS B., KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C.,  
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
 RA NOONE D., O'REILLY M., OGAWA K., OGIMARA A., OUDSGA B., PARK S.H.,  
 RA PARRO V., POHL T.M., PORTELETT D., PORKOLLIK S., PRESCOTT A.M.,  
 RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
 RA RIEGER M., RIVOLTA C., ROCHE E., ROCHE B., ROSE M., SADAIE Y.,  
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,  
 RA SEKIGUCHI J., SEKOWSKA A., SERO S.J., SERROR P., SHIN B.S., SOLDI B.,  
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
 RA TAKEUCHI M., TANAKASHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
 RA VIARI A., WAMBUIT R., WEDLER E., WEDLER H., WEITZENEGGER T.,  
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,  
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;  
 RT "The complete genome sequence of the gram-positive bacterium Bacillus  
 subtilis";  
 RL Nature 390:249-256(1997).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC6633;  
 RX MEDLINE: 93175869.  
 RA KLEIN C., KALETTA C., ENTIAN K.D.;  
 RT "Biosynthesis of the lantibiotic subtilin is regulated by a histidine  
 RT kinase/response regulator system";  
 RL Appl. Environ. Microbiol. 59:296-303(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC6633;  
 RA KLEIN C.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC6633;  
 RA KLEIN C.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z99121; CAB15371.1; -;  
 DR EMBL: U09819; AAB91590.1; -;  
 DR PFAM: PF01381; HTH\_3; 1.  
 SQ SEQUENCE 135 AA; 15445 MW; BF5F1F8F CRC32;

Query Match 60.3%; Score 35; DB 2; Length 135;  
 Best Local Similarity 66.7%; Pred. No. 24;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTKCKFLKK 9  
 :||| |:  
 DB 81 ETRCKLLEK 89

RESULT 15  
 Q14660

ID Q14660 PRELIMINARY; PRT; 2713 AA.  
 AC Q14660;  
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
 DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)  
 DE INOSITOL 1,4,5-TRISPHOSPHATE RECEPTOR TYPE 1.  
 GN ITPR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-T LYMPHOCYTE;  
 RX MEDLINE: 95155356.  
 RA HARNICK D.J., JAYARAMAN T., MA Y., MULIERI P., GO L.O., MARKS A.R.;  
 RT "The human type 1 inositol 1,4,5-trisphosphate receptor from T  
 RT lymphocytes. Structure, localization, and tyrosine phosphorylation.";  
 RL J. Biol. Chem. 270:2833-2840(1995).  
 DR EMBL: L38019; AAB04947.1; -;  
 DR PFAM: PF01365; RYDR\_ITPR; 1.  
 SQ SEQUENCE 2713 AA; 308697 MW; A16B46FB CRC32;

Query Match 60.3%; Score 35; DB 4; Length 2713;  
 Best Local Similarity 60.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTKCKFLKK 10  
 | : |||:  
 DB 1319 KAEGKFLKK 1328

Search completed: September 7, 1999, 22:47:42

Job time: 7962 sec

**THIS PAGE BLANK (USPTO)**

---

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:03 ; Search time 135.78 Seconds  
(without alignments)  
1.047 Million cell updates/sec

Title: US-09-124-280A-32  
Perfect score: 30  
Sequence: 1 KFLKKT 6  
Scoring table: BLOSUM62  
Searched: 188963 seqs, 23686106 residues  
Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	6	1 R33527	Peptide for treati
2	30	100.0	10	1 R33528	Peptide for treati
3	30	100.0	6	1 R39289	Endotoxin lipid A
4	30	100.0	10	1 R39293	Endotoxin lipid A
5	30	100.0	10	1 R57882	Lipid A minimum bi
6	30	100.0	7	1 R71777	Peptide neutralisi
7	30	100.0	6	1 W21620	Antibiotic potenti
8	30	100.0	10	1 W21624	Antibiotic potenti
9	30	100.0	7	1 W21594	Antibiotic potenti
10	30	100.0	309	1 W65049	Tsuga pinosresinol
11	30	100.0	10	1 W62433	Human neutrophil 9
12	27	90.0	503	1 W59164	L. lactis M61316 9
13	26	86.7	448	1 P93343	Gene encoding the
14	26	86.7	453	1 R40856	43kd regression as
15	26	86.7	432	1 R67582	Cancer metastasis
16	26	86.7	521	1 W50007	Human protease FMH
17	26	86.7	535	1 W77299	Amino acid sequenc
18	25	83.3	733	1 R06110	Sequence of cyclom
19	25	83.3	933	1 R08132	Beta-galactosidase
20	25	83.3	15	1 R10999	Magainin I or II a
21	25	83.3	16	1 R10639	Magainin I or II a
22	25	83.3	15	1 R10640	Magainin I or II a
23	25	83.3	22	1 R10657	Magainin II analog
24	25	83.3	22	1 R10663	Magainin I or II a
25	25	83.3	712	1 R10051	Cyclomaltodextrin
26	25	83.3	713	1 R10052	Cyclomaltodextrin
27	25	83.3	713	1 R12743	Enzyme with starch
28	25	83.3	22	1 R15795	Magainin II analog
29	25	83.3	793	1 R20743	Murine receptor ty
30	25	83.3	802	1 R20744	Human receptor-ty
31	25	83.3	351	1 R31599	Chicken nov protei
32	25	83.3	22	1 R35387	Amphiphilic peptid
33	25	83.3	22	1 R35389	Amphiphilic peptid
34	25	83.3	22	1 R35390	Amphiphilic peptid
35	25	83.3	7	1 R33525	Peptide for treati
36	25	83.3	10	1 R33526	Peptide for treati
37	25	83.3	11	1 R33531	Peptide for treati
38	25	83.3	16	1 R36380	Amphiphilic ion ch
39	25	83.3	22	1 R39080	Biologically activ
40	25	83.3	7	1 R39287	Endotoxin lipid A
41	25	83.3	10	1 R39288	Endotoxin lipid A
42	25	83.3	11	1 R39292	Endotoxin lipid A
43	25	83.3	22	1 R50566	Amphiphilic pepti

ALIGNMENTS

RESULT 1

R33527 ID R33527 standard; peptide; 6 AA.  
AC R33527;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 5; Page 32; 39pp; English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 6 AA;

Query Match 100.0%; Score 30; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKKT 6  
Db 1 KFLKKT 6

RESULT 2

R33528 ID R33528 standard; peptide; 10 AA.  
AC R33528;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PS Claim 9; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 30; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKKT 6  
 Db 5 KFLKKT 10

RESULT 3  
 R39289 ID R39289 standard; peptide; 6 AA.  
 AC R39289;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M.  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 5; Page 32; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 6 AA;

Query Match 100.0%; Score 30; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKKT 6  
 Db 1 KFLKKT 6

RESULT 4  
 R39293 ID R39293 standard; peptide; 10 AA.  
 AC R39293;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M.  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 9; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 30; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKKT 6  
 Db 5 KFLKKT 10

RESULT 5  
 R57882 ID R57882 standard; peptide; 10 AA.  
 AC R57882;  
 DT 28-MAR-1995 (first entry)  
 DE Lipid A minimum binding site.  
 KW Binding site; CDR; complementarity determining region; immunoglobulin;  
 KW heavy; light; primer extension; PCR; amplify; fibronectin; vitronectin;  
 KW RGD-dependent; integrin ligand; von Willebrand factor; EBV; gp350/220;  
 KW envelope glycoprotein; HIV; gp120; reovirus; hemagglutinin; insulin;  
 KW cellular receptor; CR2; CD4; hormone; thyroid stimulating hormone; TSH;  
 KW transferrin; apolipoprotein; apo E; apo A1; MHC; class I; class II;  
 KW non-RGD-dependent; vitronectin receptor; alpha-v, beta-3; modulation;  
 KW anti-gp11b/IIIa; monoclonal antibody; MAB; platelet adhesion; cancer;  
 KW coagulation; inflammation; anti-vitronectin; tumour cell adhesion;  
 KW migration.  
 OS Homo sapiens.  
 PN WO9418221-A.  
 PD 18-AUG-1994.  
 PF 02-FEB-1994; U01258.  
 PR 02-FEB-1993; US-012566.  
 PR 28-JUN-1993; US-084542.

PA (SCRI ) SCRIPPS RES INST.  
 PI Barbas CF, Lerner RA;  
 DR WFI: 94-279675/34.  
 PT Production of binding sites within CDR regions of immunoglobulins  
 PT - displayed on the surface of filamentous phage particles, for  
 PS inhibiting platelet aggregation and vitronectin binding  
 PS Disclosure: Page 27: 207pp; English.  
 CC The sequences given in R57837-84 are binding sites which were used in  
 CC the method of the invention for producing a polypeptide having a  
 CC binding site capable of binding a preselected agent. Nucleotide  
 CC sequences encoding these binding site peptides were introduced into  
 CC a CDR region of a nucleic acid encoding an immunoglobulin heavy (H)  
 CC or light (L) chain, by amplifying the CDR region by primer extension.  
 CC Preferred binding sites are derived from the RGD-dependent integrin  
 CC ligands, eg. fibronectin, vitronectin, von Willebrand factor, from  
 CC the envelope glycoprotein from viruses such as HIV gp120, EBV gp350/  
 CC 220, reovirus hemagglutinin, from cellular receptors such as CR2 or  
 CC CD4, from protein hormones such as thyroid stimulating hormone (TSH),  
 CC insulin, transferrin, from apolipoproteins such as apo E and apo AI,  
 CC from immunoglobulin CDRs and from MHC class I or II proteins. Non-RGD-  
 CC dependent integrin binding sites were selected for the affinity to bind  
 CC vitronectin receptor alpha-v, beta-3. An anti-gp11b/IIb monoclonal  
 CC antibody (Mab) produced in this way can be used to modulate platelet  
 CC adhesion in the treatment of coagulation and some inflammatory responses.  
 CC An anti-vitronectin Mab can be used in the treatment of cancer by  
 CC blocking tumour cell adhesion and migration. This sequence represents  
 CC the lipid A binding site.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 30; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 |||||  
 DB 5 KFLKKT 10

## RESULT 6

R71777  
 ID R71777 standard; peptide; 7 AA.  
 AC R71777; 1995 (first entry).  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WFI: 95-075190/10.  
 PT New peptides for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 8; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)<sub>n</sub>, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)<sub>m</sub>, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)<sub>p</sub>, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 30; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 |||||  
 DB 1 KFLKKT 6

## RESULT 7

W21620  
 ID W21620 standard; peptide; 6 AA.  
 AC W21620; 1997 (first entry)  
 DE Antibiotic potentiating peptide #32.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 03-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WFI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 36; Page 27; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 6 AA;

Query Match 100.0%; Score 30; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 |||||  
 DB 1 KFLKKT 6

## RESULT 8

W21624  
 ID W21624 standard; peptide; 10 AA.  
 AC W21624; 1997 (first entry)  
 DE Antibiotic potentiating peptide #36.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 03-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WFI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 40; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of

CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 30; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
 |||||  
 Db 5 KFLKKT 10

## RESULT 9

W21594  
 ID W21594 standard; peptide; 7 AA.  
 AC W21594;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #6.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 10: Page 24: 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 30; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
 |||||  
 Db 1 KFLKKT 6

## RESULT 10

W65049  
 ID W65049 standard; Protein; 309 AA.  
 AC W65049;  
 DT 28-SEP-1998 (first entry)  
 DE Tsuga pinosresinol-lariciresinol reductase PLR-Th2.  
 KW Pinosresinol-lariciresinol reductase; dirigent protein; lignan;  
 KW transgenic plant; anticancer; cytostatic; antiviral; virucide;  
 KW antibiotic; antioxidant; antifeedant.  
 OS Tsuga heterophylla.  
 PN WO9820113-A1.  
 PD 14-MAY-1998.  
 PF 07-NOV-1997; U20391.  
 PR 31-JUL-1997; US-054380.  
 PR 08-NOV-1996; US-030522.  
 PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PI Davin LB, Dinkovakostova AT, Fujita M, Gang DR,

PI Lewis NG, Sarkanen S;  
 DR WPI: 98-286929/25.  
 DR N-PSDB: V35187.  
 PT New plant-derived dirigent proteins and pinosresinol-lariciresinol  
 PT reductases - and related nucleic acid, vectors and transformants,  
 PT used for stereospecific production of lignans, useful e.g. as  
 PT anticancer or antiviral agents  
 PS Claim 17: Page 138-139; 148pp; English.  
 CC PLR-Th2 comprises a (+)-pinosresinol/(+)-lariciresinol reductase  
 CC (P/LR) of Tsuga heterophylla. P/LR enzymes catalyse the conversion  
 CC of pinosresinol to lariciresinol and then to secoisolariciresinol.  
 CC 2 isoforms (see W65048-49) of T. heterophylla P/LR have been  
 CC identified. Further P/LR enzymes have been obtained from Forsythia  
 CC intermedia (see W65038-43) and Thuja plicata (see W65044-47). The  
 CC isolation of cDNAs encoding P/LR (see V35175-80 and V35182-87) and  
 CC dirigent proteins (see V35160-71) permits the development of an  
 CC efficient expression system for these enzymes, provides useful  
 CC tools for examining the developmental regulation of lignan  
 CC biosynthesis and permits the isolation of related sequences. It  
 CC also allows the transformation of a wide range of organisms,  
 CC including plants, in order to modify lignan biosynthesis.  
 CC Optically pure lignans may have e.g. anticancer, antiviral,  
 CC antioxidant, antibiotic or antifeedant activity.  
 SQ Sequence 309 AA;

Query Match 100.0%; Score 30; DB 1; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
 |||||  
 Db 238 KFLKKT 243

## RESULT 11

W62433  
 ID W62433 standard; peptide; 10 AA.  
 AC W62433;  
 DT 01-OCT-1998 (first entry)  
 DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:48.  
 KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;  
 KW bactericidal permeability increasing factor; B/PI; neutralising;  
 KW human neutrophil granule bactericidal protein.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN US5786324-A.  
 PD 28-JUL-1998.  
 PF 24-MAR-1994; 218026.  
 PR 24-MAR-1994; US-218026.  
 PA (MINU ) UNIV MINNESOTA.  
 PI Gray B, Haseman JR, Mayo K;  
 DR WPI: 98-436578/37.  
 PT Bactericidal and endotoxin-neutralising peptides - used in treating  
 PT e.g. Pseudomonas species infection and in protectively coating  
 PT prosthetic devices  
 PS Example 1: Column 18: 46pp; English.  
 CC The present invention describes bactericidal peptides (BP): (a) for  
 CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)  
 CC with both endotoxin neutralising activity and bactericidal activity for  
 CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising  
 CC activity and bactericidal activity, the peptide having at least 1 amino  
 CC acid sequence of a peptide with endotoxin neutralising activity combined  
 CC with at least 1 different amino acid sequence of a peptide that has  
 CC bactericidal activity, and (e) a prosthetic device which has a  
 CC sufficient amount of BP attached to the surface to inhibit bacterial  
 CC growth. The peptides of the invention are used in treating bacterial  
 CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7  
 CC 10-9 M, and Escherichia coli. The peptides are also used to treat  
 CC endotoxin shock. The present sequence represents a peptide derived  
 CC from human neutrophil granule bactericidal protein from an example of  
 CC the present invention.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 30; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 |||||  
 Db 5 KFLKKT 10

RESULT 12  
 W59164  
 ID W59164 standard; Protein; 503 AA.  
 AC W59164;  
 DT 13-AUG-1998 (first entry)  
 DE L. lactis MG1316 gadC protein from PstI/EcoRI fragment.  
 KW Salt-inducible promoter; lactic acid; food industry; food-grade inducer;  
 KW fermentation processes; cheese production; rhnB; gadC; orfX; rggL.  
 OS Lactococcus lactis.  
 PN WO9810080-A1.  
 PD 12-MAR-1998.  
 PF 20-AUG-1997; E04755.  
 PR 13-MAR-1997; EP-200744.  
 PR 05-SEP-1996; EP-202444.  
 PA (UNIL ) UNILEVER NV.  
 PA (UNIL ) UNILEVER PLC.  
 PI Kok J, Ledebor AM, Sanders JW, Venema G;  
 DR WPI: 98-193629/17.  
 DR N-PSDB; V11905.  
 PT Salt-inducible promoter - derived from lactic acid bacteria, used  
 PT for the production of polypeptides in food  
 PS Example 6; Fig 29; 111pp; English.  
 CC This protein sequence is the gadC protein (also known as orfX) found in  
 CC a PstI-EcoRI fragment from the NS3 locus of the lactic acid bacterium  
 CC Lactococcus lactis. This fragment also contains the C-terminus of rhnB,  
 CC full length gadR protein (also known as rggL), gadB and the C-terminus of  
 CC an unknown reading frame which reads in the reverse orientation. This  
 CC fragment is also found to act as a salt-inducible promoter (SIP). Using  
 CC this SIP, salt can be used as a food-grade inducer in food fermentation  
 CC processes, e.g. in the production of cheese, dressings, water-containing  
 CC spreads, sausages, or sour dough.  
 SQ Sequence 503 AA;

Query Match 90.0%; Score 27; DB 1; Length 503;  
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 |||||  
 Db 320 KFLKKT 325

RESULT 13  
 P93343  
 ID P93343 standard; protein; 448 AA.  
 AC P93343;  
 DT 01-AUG-1990 (first entry)  
 DE Gene encoding the 43 kd regression-associated antigen (RA Ag) of  
 DE Mycoplasma hyorhinis  
 KW Tumour regression-associated antigens (RA Ag); in vivo imaging;  
 KW therapy monitoring; cancer therapy; Mycoplasma hyorhinis.  
 OS Mycoplasma hyorhinis.  
 FH Location/Qualifiers  
 FT peptide 1..25  
 FT protein 26..448  
 PN EP-308265-A.  
 PD 22-MAR-1989.  
 PF 15-SEP-1988; 308625.  
 PR 11-DEC-1987; US-131815, US-097910.  
 PA (ITGE-) Int Genetic Eng Inc.  
 PI Fareed GC, Sen A, Ghosh-Dastidar P, Liu A, Lee JH;

DR WPI: 89-087638/12.  
 DR N-PSDB; N90684.  
 PT Tumour regression-associated antigens and antibodies -  
 PT used in diagnostic tests, monitoring course of therapy and for  
 PT therapy in cancer patients  
 PS Disclosure: 56pp; English.  
 CC Regression-associated antigens may be used in diagnostic tests, eg in  
 CC vivo imaging, for monitoring the course of therapy or for therapeutic  
 CC purposes, eg active immunisation protocols in cancer patients or drug  
 CC delivery systems by binding the drug to monoclonal or monospecific  
 CC polyclonal Ab showing specific immunoreactivity with the Ag.  
 SQ Sequence 448 AA;

Query Match 86.7%; Score 26; DB 1; Length 448;  
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 |||||  
 Db 137 KFLKQT 142

RESULT 14  
 R40856  
 ID R40856 standard; Protein; 453 AA.  
 AC R40856;  
 DT 07-MAR-1994 (first entry)  
 DE 43kd regression associated antigen.  
 KW Regression associated antigen; tumour; immunotherapy;  
 KW anti-idiotypic antibodies; antibodies; tumour regression.  
 OS Mycoplasma hyorhinis.  
 FH Location/Qualifiers  
 FT misc\_difference 80  
 FT /note= "Tryptophan encoded by TGA, normal in  
 FT Mycoplasma hyorhinis."  
 FT misc\_difference 124  
 FT /note= "Tryptophan encoded by TGA, normal in  
 FT Mycoplasma hyorhinis."  
 FT misc\_difference 165  
 FT /note= "Tryptophan encoded by TGA, normal in  
 FT Mycoplasma hyorhinis."  
 FT misc\_difference 344  
 FT /note= "Tryptophan encoded by TGA, normal in  
 FT Mycoplasma hyorhinis."  
 FT US5242823-A.  
 PD 07-SEP-1993.  
 PF 07-MAR-1986; 837494.  
 PR 07-MAR-1986; US-837494.  
 PR 16-SEP-1987; US-097910.  
 PR 11-DEC-1987; US-131815.  
 PR 04-JAN-1988; US-138923.  
 PR 16-MAR-1990; US-474730.  
 PR 02-OCT-1992; US-956546.  
 PA (ITGE-) INT GENETIC ENG INC.  
 PI Fareed GC, Ghosh-dastidar P, Jar-how L, Sen A;  
 DR WPI: 93-295229/37.  
 DR N-PSDB; Q47816.  
 PT DNA encoding a regression-associated antigen from M. hyorhinis -  
 PT is used to obtain prods. for diagnosis, localisation and therapy  
 PT of tumours  
 PS Disclosure; Figure 3; 40pp; English.  
 CC Regression associated antigens (RAA's) are identified in material  
 CC from neoplastic cells by their immunological reactivity with  
 CC regression associated antibodies from the serum of patients  
 CC diagnosed as undergoing regression of a tumour. RAA's can be used  
 CC for tumour immunotherapy and for producing and purifying antibodies  
 CC which can be used for tumour diagnosis, localisation and therapy.  
 CC The antibodies can also be used for the production of  
 CC anti-idiotypic antibodies which can also be used in immunotherapy.  
 SQ Sequence 453 AA;

Query Match 86.7%; Score 26; DB 1; Length 453;  
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 |||||  
 Db 137 KFLKQT 142

## RESULT 15

R67582  
 ID R67582 standard; Protein: 432 AA.  
 AC R67582;  
 DT 06-OCT-1995 (first entry)  
 DE Cancer metastasis protein.  
 KW Cancer metastasis; clinical tests; detection.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT misc\_difference 55 /note= "corresponding codon TGA"  
 FT misc\_difference 99 /note= "corresponding codon TGA"  
 FT misc\_difference 140 /note= "corresponding codon TGA"  
 FT misc\_difference 319 /note= "corresponding codon TGA"  
 FT misc\_difference 420 /note= "corresponding codon TGA"  
 FT misc\_difference 424 /note= "corresponding codon TGA"  
 FT J06319559-A.  
 FN 22-NOV-1994.  
 PD 12-MAY-1993; 144165.  
 PR 12-MAY-1993; JP-144165  
 PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 DR WPI: 95-040317/06.  
 DR N-PSDB: Q79124.  
 PT DNA encoding polypeptide involved in cancer metastasis - useful  
 PT for the study of the mechanism of metastasis and in clinical  
 PT tests  
 PS Claim 1; Page 2; 13pp; Japanese.  
 CC Q79124 encodes R67582 a protein involved in cancer metastasis,  
 CC which may be used for the study of metastatic mechanisms, and for  
 CC clinical tests to determine the presence or absence of cancer  
 CC metastasis.  
 SQ Sequence 432 AA;

Query Match 86.7%; Score 26; DB 1; Length 432;  
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 |||||  
 Db 112 KFLKQT 117

Search completed: September 7, 1999, 22:50:04  
 Job time: 7885 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:49 ; Search time 82.37 Seconds  
(without alignments)  
0.719 Million cell updates/sec

Title: US-09-124-280A-32

Perfect score: 30

Sequence: 1 KFLKKT 6

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/PTUS9\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	6	1	US-08-049-871-3
2	30	100.0	10	1	US-08-049-871-7
3	30	100.0	6	1	US-07-819-893-3
4	30	100.0	10	1	US-07-819-893-7
5	30	100.0	6	1	US-08-280-397-3
6	30	100.0	10	1	US-08-280-397-7
7	30	100.0	7	1	US-08-087-830E-6
8	30	100.0	6	1	US-08-097-830E-30
9	30	100.0	10	2	US-08-218-026-48
10	30	100.0	10	2	US-08-653-632-48
11	30	100.0	7	2	US-08-456-112B-6
12	30	100.0	6	2	US-08-456-112B-32
13	30	100.0	10	2	US-08-456-112B-36
14	30	100.0	7	3	PCT-US94-01234-46
15	25	83.3	10	1	US-08-049-871-1
16	25	83.3	10	1	US-08-049-871-2
17	25	83.3	11	1	US-08-049-871-6
18	25	83.3	7	1	US-07-819-893-1
19	25	83.3	10	1	US-07-819-893-2
20	25	83.3	11	1	US-07-819-893-6
21	25	83.3	21	1	US-07-965-663A-1
22	25	83.3	21	1	US-07-965-663A-2
23	25	83.3	21	1	US-07-965-663A-3
24	25	83.3	22	1	US-07-965-663A-6
25	25	83.3	22	1	US-07-965-663A-7
26	25	83.3	23	1	US-07-965-663A-8
27	25	83.3	23	1	US-07-965-663A-9
28	25	83.3	23	1	US-07-965-663A-10
29	25	83.3	23	1	US-07-965-663A-11
30	25	83.3	24	1	US-07-965-663A-12
31	25	83.3	25	1	US-07-965-663A-13
32	25	83.3	25	1	US-07-965-663A-14
33	25	83.3	26	1	US-07-965-663A-15
34	25	83.3	23	1	US-07-965-663A-16
35	25	83.3	24	1	US-07-965-663A-18
36	25	83.3	22	1	US-07-965-663A-19
37	25	83.3	22	1	US-07-965-663A-20
38	25	83.3	22	1	US-07-965-663A-21
39	25	83.3	21	1	US-07-965-663A-22

Sequence 23, Appl  
Sequence 1, Appl1  
Sequence 3, Appl1  
Sequence 6, Appl1  
Sequence 4, Appl1  
Sequence 51, Appl

ALIGNMENTS

RESULT 1

US-08-049-871-3  
; Sequence 3, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: PORRO, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-049-871-3

Query Match 100.0%; Score 30; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKKT 6  
Db 1 KFLKKT 6

RESULT 2

US-08-049-871-7  
; Sequence 7, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: PORRO, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the

;  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-049-871-7

Query Match 100.0%; Score 30; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1-2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
Db 5 KFLKKT 10

RESULT 3  
US-07-819-893-3  
; Sequence 3, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514

;  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
; US-07-819-893-3

Query Match 100.0%; Score 30; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
Db 1 KFLKKT 6

RESULT 4  
US-07-819-893-7  
; Sequence 7, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; US-07-819-893-7

Query Match 100.0%; Score 30; DB 1; Length 10;

```
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLAKT 6
Db 5 KFLAKT 10

RESULT 5
US-08-280-397-3
; Sequence 3, Application US/08280397
; Patent No. 5589459
; GENERAL INFORMATION:
; APPLICANT: Porto, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,397
; FILING DATE: 07/26/94
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/819,893
; FILING DATE: 01/16/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8989
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-08-280-397-3

Query Match 100.0%; Score 30; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLAKT 6
Db 1 KFLAKT 6

RESULT 6
US-08-280-397-7
; Sequence 7, Application US/08280397
; Patent No. 5589459
; GENERAL INFORMATION:
; APPLICANT: Porto, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.

Best Local Similarity 100.0%; Score 30; DB 1; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLAKT 6
Db 5 KFLAKT 10

RESULT 7
US-08-097-830E-6
; Sequence 6, Application US/08097830E
; Patent No. 5652211
; GENERAL INFORMATION:
; APPLICANT: Porto, Massimo
; TITLE OF INVENTION: Peptides For Neutralizing The
; TITLE OF INVENTION: Toxicity of Lipid A
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
```

REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-6

Query Match 100.0%; Score 30; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
Db 1 KFLKKT 6

RESULT 8  
US-08-097-830E-30  
Sequence 30, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TITLE OF INVENTION: Toxicity of Lipid A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-097-830E-30

Query Match 100.0%; Score 30; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
Db 1 KFLKKT 6

## RESULT 9

US-08-218-026-48  
Sequence 48, Application US/08218026  
Patent No. 5786324  
GENERAL INFORMATION:

APPLICANT: Gray, Beulah  
APPLICANT: Haseman, Judith R.

TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
Activity and Endotoxin Neutralizing Activity for Gram  
TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5786324west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/218,026  
FILING DATE: 24-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalchuk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 600.286US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-218-026-48

Query Match 100.0%; Score 30; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
Db 5 KFLKKT 10

## RESULT 10

US-08-653-632-48  
Sequence 48, Application US/08653632  
Patent No. 5830860  
GENERAL INFORMATION:

APPLICANT: GRAY, Beulah  
APPLICANT: HASEMAN, Judith R.

TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 5830860west Center, 90 South Seventh St  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

```
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/653,632
; FILING DATE: 24-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/218026
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Denise M
; REGISTRATION NUMBER: 35,924
; REFERENCE/DOCKET NUMBER: 600.286US11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/371-5268
; TELEFAX: 612/332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-653-632-48

Query Match 100.0%; Score 30; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6
Db 5 KFLKKT 10

RESULT 11
US-08-456-112B-6
; Sequence 6, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: POIRO, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; US-08-456-112B-32

Query Match 100.0%; Score 30; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6
Db 1 KFLKKT 6

RESULT 13
US-08-456-112B-36
; Sequence 36, Application US/08456112B
```

```
;
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; US-08-456-112B-6

Query Match 100.0%; Score 30; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6
Db 1 KFLKKT 6

RESULT 12
US-08-456-112B-32
; Sequence 32, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: POIRO, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; US-08-456-112B-32

Query Match 100.0%; Score 30; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6
Db 1 KFLKKT 6

RESULT 13
US-08-456-112B-36
; Sequence 36, Application US/08456112B
```

```
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; US-08-456-112B-36

Query Match 100.0%; Score 30; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6
Db 5 KFLKKT 10

RESULT 14
PCT-US94-01234-46
; Sequence 46, Application PC/TUS9401234
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE
; TITLE OF INVENTION: BINDING SITES
; NUMBER OF SEQUENCES: 76
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01234
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/084,542
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
PCT-US94-01234-46

Query Match 100.0%; Score 30; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6
Db 5 KFLKKT 10

RESULT 15
US-08-049-871-1
; Sequence 1, Application US/08049871
; Patent No. 5358933
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,871
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/658,744
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; US-08-049-871-1

Query Match 83.3%; Score 25; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKK 5
Db 2 KFLKK 6
```

Search completed: September 7, 1999, 23:07:49  
Job time: 1747 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:46 ; Search time 105.14 Seconds  
(without alignments)  
2.286 Million cell updates/sec

Title: US-09-124-280A-32  
Perfect score: 30  
Sequence: 1 KFLKKT 6

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	145	2 C49683	pts 3'-region prot
2	30	100.0	764	2 T01493	high affinity pota
3	28	93.3	307	1 WMNV38	PE-38 protein - Or
4	28	93.3	232	2 A64366	hypothetical prote
5	28	93.3	743	2 T00634	hypothetical prote
6	27	90.0	410	2 C39088	contrapsin precurs
7	27	90.0	202	2 B71509	probable phosphata
8	27	90.0	507	2 T00871	probable 7-ethoxyc
9	27	90.0	1435	2 S69632	regulatory protein
10	26	86.7	650	1 ISYMBP	DNA topoisomerase
11	26	86.7	1094	2 S22573	DNA-directed DNA p
12	26	86.7	813	2 D70176	ATP-dependent prot
13	26	86.7	148	2 A39106	ribosomal protein
14	26	86.7	1058	2 S63460	apolipoprotein B -
15	26	86.7	173	2 S44283	pail repressor hom
16	26	86.7	305	2 B71643	probable integrase
17	26	86.7	659	2 S65976	hypothetical prote
18	26	86.7	1447	2 S63669	UDPglucose-glycop
19	26	86.7	250	2 E70452	hypothetical prote
20	25	83.3	713	1 ALBSG1	cyclomaltodextrin
21	25	83.3	712	1 ALBSG3	cyclomaltodextrin
22	25	83.3	713	1 ALBSG7	cyclomaltodextrin
23	25	83.3	1192	1 ISXPAS	DNA topoisomerase
24	25	83.3	84	1 R3EC17	ribosomal protein
25	25	83.3	1191	1 A44051	DNA-binding protei
26	25	83.3	321	1 A43681	immediate-early pr
27	25	83.3	493	2 JCA372	IMP dehydrogenase
28	25	83.3	499	2 A30151	dihydrolipoamide d
29	25	83.3	463	2 B36936	nitrogenase (EC 1
30	25	83.3	1161	2 B70172	DNA polymerase III
31	25	83.3	2282	2 S16664	large protein L -
32	25	83.3	713	2 S09196	cyclomaltodextrin
33	25	83.3	713	2 A58800	cyclomaltodextrin
34	25	83.3	686	2 E64946	oligopeptidase B (
35	25	83.3	235	2 E42696	thrombin (EC 3.4.2
36	25	83.3	248	2 S58093	probable multicata
37	25	83.3	1191	2 S27329	DNA topoisomerase
38	25	83.3	1301	2 A41622	protein-tyrosine-p
39	25	83.3	802	2 A36065	protein-tyrosine-p

40 25 83.3 796 2 JC1285 protein-tyrosine-p  
41 25 83.3 829 2 A47373 protein-tyrosine-p  
42 25 83.3 700 2 S12053 protein-tyrosine-p  
43 25 83.3 109 2 B61180 protein-tyrosine-p  
44 25 83.3 699 2 JC6132 protein-tyrosine-p  
45 25 83.3 4869 2 S66572 ryanodine receptor

#### ALIGNMENTS

##### RESULT 1

C49683

pts 3'-region protein 2 - Mycoplasma capricolium (SGC3) (fragment)

C:Species: Mycoplasma capricolium

C>Date: 26-May-1995 #sequence\_revision 26-May-1995 #text\_change 09-Sep-1997

C:Accession: C49683

R:Zhu, P.P.; Reizer, J.; Reizer, A.; Peterkofsky, A.

J. Biol. Chem. 268, 26531-26540, 1993

A:Title: Unique monocistronic operon (ptsH) in Mycoplasma capricolium encoding the photoactivation of ptsH.

A:Reference number: A49683; MUID:94075343

A:Accession: C49683

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-145 <ZHU>

A:Cross-references: GB:L22432; NID:g435095; PID:g435099

C:Genetics:

A:Genetic code: SGC3

Query Match 100.0%; Score 30; DB 2; Length 145;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6

DB 17 KFLKKT 22

##### RESULT 2

T01493

high affinity potassium transporter protein homolog F1707.17 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 12-Feb-1999

C:Accession: T01493

R:Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel submitted to the EMBL Data Library, June 1998

A:Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.

A:Reference number: Z14334

A:Accession: T01493

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-764 <VYS>

A:Cross-references: EMBL:AC003671; NID:g2833627; PID:g3176686

C:Genetics:

A:Map position: 1

A:Introns: 12/3; 178/1; 282/3; 409/1

A:Note: F1707.17

Query Match 100.0%; Score 30; DB 2; Length 764;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6

DB 243 KFLKKT 248

##### RESULT 3

WMNV38

PE-38 protein - Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus  
 C:Species: Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus, OPMNPV  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 13-Sep-1997  
 C:Accession: A42191  
 R:Theilmann, D.A.; Stewart, S.  
 Virology 187, 97-106, 1992  
 A:Title: Tandemly repeated sequence at the 3' end of the IE-2 gene of the baculovirus Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus  
 A:Reference number: A42191; MUID:92142537  
 A:Accession: A42191  
 A:Molecule type: DNA  
 A:Residues: 1-307 <THE>  
 A:Cross-references: GB:M83827; NID:g332540; PID:g332542  
 C:Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus PE-38 protein  
 C:Keywords: DNA binding; early protein

Query Match 93.3%; Score 28; DB 1; Length 307;  
 Best Local Similarity 83.3%; Pred. No. 34;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 ||:|||  
 Db 110 KFIKKT 115

RESULT 4  
 A64366  
 hypothetical protein MJ0529 - Methanococcus jannaschii  
 C:Species: Methanococcus jannaschii  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 10-Oct-1997  
 C:Accession: A64366  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A:Reference number: A64300; MUID:96337999  
 A:Accession: A64366  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-232 <BUL>  
 A:Cross-references: GB:U67502; GB:I77117; NID:gi591223; PID:gi591232; TIGR:MJ0529; PID:g332542  
 C:Genetics:  
 A:Map position: REV466217-465519

Query Match 93.3%; Score 28; DB 2; Length 232;  
 Best Local Similarity 83.3%; Pred. No. 26;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLAKT 6  
 ||:|||  
 Db 197 KFMKKT 202

RESULT 5  
 T00634  
 hypothetical protein\_HDJ0897G10.1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 01-Feb-1999  
 C:Accession: T00634  
 R:Kalicki, J.; Elliott, G.  
 submitted to the EMBL Data Library, January 1998  
 A:Description: The sequence of Homo sapiens PAC clone DJ0897G10.  
 A:Reference number: Z14194  
 A:Accession: T00634  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-743 <KAL>  
 A:Cross-references: EMBL:AC004082; NID:g2822160; PID:g2822161  
 C:Genetics:  
 A:Map position: 7q11.23-q21.1

A; Introns: 36/2; 59/1; 130/3; 172/2; 203/3; 227/3; 288/1; 353/1; 414/1; 461/2; 486/1;  
 A; Note: H\_DJ0897G10.1

Query Match 93.3%; Score 28; DB 2; Length 743;  
 Best Local Similarity 83.3%; Pred. No. 82;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 ||:|||  
 Db 706 KFMKKT 711

RESULT 6  
 C39088  
 contrapsin precursor - guinea pig  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 27-Nov-1991 #sequence\_revision 27-Nov-1991 #text\_change 12-Apr-1996  
 C:Accession: C39088  
 R:Suzuki, Y.; Yoshida, K.; Honda, E.; Sinozawa, H.  
 J. Biol. Chem. 266, 928-932, 1991  
 A:Title: Molecular cloning and sequence analysis of cDNAs coding for guinea pig alpha  
 A:Reference number: A39088; MUID:91093294  
 A:Accession: C39088  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-410 <SUZ>  
 A:Cross-references: GB:M38573  
 C:Superfamily: antithrombin III

Query Match 90.0%; Score 27; DB 2; Length 410;  
 Best Local Similarity 83.3%; Pred. No. 77;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 ||:|||  
 Db 290 KFLKKT 295

RESULT 7  
 B71509  
 probable phosphatase/kinase - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 21-Nov-1998  
 C:Accession: B71509  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
 Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t  
 A:Reference number: A71570; MUID:9900809  
 A:Accession: B71509  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-202 <ARN>  
 A:Cross-references: GB:AE001322; GB:AE001273; NID:g3328916; PID:g3328928  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: yacE

Query Match 90.0%; Score 27; DB 2; Length 202;  
 Best Local Similarity 83.3%; Pred. No. 38;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 ||:|||  
 Db 145 RFLKKT 150

RESULT 8  
 T00871  
 probable 7-ethoxycoumarin O-deethylase F17K2.11 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)



C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 12-Feb-1999  
C;Accession: T00871  
R;Roundley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, March 1998  
A;Description: Arabidopsis thaliana chromosome II BAC F17K2 genomic sequence.  
A;Reference number: 214207  
A;Accession: T00871  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-507 <ROU>  
A;Cross-references: EMBL:AC003680; NID:g2979540; PID:g2979550  
C;Genetics:  
A;Map position: II  
A;Introns: 123/2; 300/3; 393/3  
A;Note: F17K2.11

Query Match 90.08; Score 27; DB 2; Length 507;  
Best Local Similarity 83.3%; Pred. No. 95;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
:|||||  
Db 123 RFLKKT 128

RESULT 9  
S69632  
N;Alternate names: protein YDR464w  
C;Species: Saccharomyces cerevisiae  
C;Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 17-Mar-1999  
C;Accession: S69632; S47864  
R;Dietrich, F.S.  
submitted to the EMBL Data Library, August 1995  
A;Description: The sequence of S. cerevisiae cosmid8 9410, 8035, 8166, and 9787.  
A;Reference number: S69553  
A;Accession: S69632  
A;Molecule type: DNA  
A;Residues: 1-1435 <DIE>  
A;Cross-references: EMBL:U33050; NID:g927726; PID:g927729; MIPS:YDR464w  
R;Maddock, J.R.; Weidenhammer, E.M.; Adams, C.C.; Lunz, R.L.; Woolford Jr., J.L. Genetics 136, 833-847, 1994  
A;Title: Extragenic suppressors of Saccharomyces cerevisiae prp4 mutations identify a new regulatory protein  
A;Reference number: S47864; MUID:94274035  
A;Accession: S47864  
A;Molecule type: DNA  
A;Residues: 1-1394, 'R' <NAD>  
A;Cross-references: EMBL:U03673; NID:g435022; PID:g435023  
C;Genetics:  
A;Gene: SGD:SPP41  
A;Cross-references: SGD:S0002872; MIPS:YDR464w  
A;Map position: 4R  
C;Keywords: nucleus

Query Match 90.08; Score 27; DB 2; Length 1435;  
Best Local Similarity 83.3%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
:|||||  
Db 1216 KYLKKT 1221

RESULT 10  
ISYMBP  
DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain B - Mycoplasma pneumoniae (SGC3)  
N;Alternate names: DNA topoisomerase type II chain B; DNA-gyrase chain B  
C;Species: Mycoplasma pneumoniae  
A;Variety: ATCC 29342  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 24-Sep-1998  
C;Accession: S11767; S62834; S73477; S14630

R;Colman, S.D.; Hu, P.C.; Bott, K.F. Mol. Microbiol. 4, 1129-1134, 1990  
A;Title: Mycoplasma pneumoniae DNA gyrase genes.  
A;Reference number: S11767; MUID:91041721  
A;Accession: S11767  
A;Molecule type: DNA  
A;Residues: 1-650 <COL>  
A;Cross-references: EMBL:X53555; NID:g44483; PID:g44484  
R;Hilbert, H.; Himmelreich, R.; Plagens, H.; Herrmann, R. Nucleic Acids Res. 24, 628-639, 1996  
A;Title: Sequence analysis of 56 kb from the genome of the bacterium Mycoplasma pneumoniae  
A;Reference number: S62797; MUID:96177562  
A;Accession: S62834  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-637 <HIL>  
A;Cross-references: EMBL:U34816; NID:g1209514; PID:g1209515  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995  
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R. Nucleic Acids Res. 24, 4420-4449, 1996  
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A;Reference number: S73327; MUID:97105885  
A;Accession: S73477  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-650 <HIM>  
A;Cross-references: EMBL:AE000016; GB:U00089; NID:g1673796; PID:g1673811  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C;Genetics:  
A;Gene: gyrB  
A;Genetic code: SGC3  
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B  
C;Keywords: antibiotic resistance; ATP; heterotetramer; isomerase

Query Match 86.7%; Score 26; DB 1; Length 650;  
Best Local Similarity 83.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
:|||||  
Db 318 RFLKKT 323

RESULT 11  
S22573  
DNA-directed DNA polymerase (EC 2.7.7.7) delta - Plasmodium falciparum  
C;Species: Plasmodium falciparum  
C;Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 20-Mar-1998  
C;Accession: S22573; A45639; S17330  
R;Ridley, R.G.; White, J.H.; McAleese, S.M.; Goman, M.; Alano, P.; de Vries, E.; Kilb Nucleic Acids Res. 19, 6731-6736, 1991  
A;Title: DNA polymerase delta: gene sequences from Plasmodium falciparum indicate the  
A;Reference number: S22573; MUID:92107655  
A;Accession: S22573  
A;Molecule type: DNA  
A;Residues: 1-1094 <RID>  
A;Cross-references: EMBL:X62423; NID:g9940; PID:g9941  
R;Fox, B.A.; Bzik, D.J. Mol. Biochem. Parasitol. 49, 289-296, 1991  
A;Title: The primary structure of Plasmodium falciparum DNA polymerase delta is similar  
A;Reference number: A45639; MUID:92131065  
A;Accession: A45639  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1094 <FOX>  
A;Cross-references: GB:M64715; NID:g160269; PID:g160270  
A;Note: sequence extracted from NCBI backbone (NCBI:77767, NCBIP:77769)  
C;Superfamily: herpesvirus DNA-directed DNA polymerase  
C;Keywords: DNA binding; DNA replication; exonuclease; nucleotidyltransferase; nucleu  
F:1049-1067/Region: zinc finger CCCC motif  
F:1049-1067/Region: zinc finger CCCC motif

Query Match 86.7%; Score 26; DB 2; Length 1094;  
Best Local Similarity 83.3%; Pred. No. 3.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|||||  
DB 995 KFLKKS 1000

## RESULT 12

D70176  
ATP-dependent proteinase La (lon-2) homolog - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 21-Aug-1998  
C:Accession: D70176  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
son, D.; Peterson, J.; Keriavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: D70176  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-813 <KLE>  
A:Cross-references: GB:AF001162; GB:AE000783; NID:q2688528; PID:q2688530; TIGR:BB0613  
A:Experimental source: strain B31  
C:Superfamily: ATP-dependent serine proteinase La  
C:Keywords: P-loop  
F:369-376/Region: nucleotide-binding motif A (P-loop)

Query Match 86.7%; Score 26; DB 2; Length 813;  
Best Local Similarity 83.3%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|||||  
DB 150 KFLKET 155

## RESULT 13

A39106  
ribosomal protein S19 e - common roundworm  
N:Alternate names: ribosomal protein ALEP-1  
C:Species: Ascaris lumbricoides (common roundworm)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Sep-1997  
C:Accession: A39106  
R:Etter, A.; Aboutanos, M.; Tobler, H.; Mueller, F.  
Proc. Natl. Acad. Sci. U.S.A. 88, 1593-1596, 1991  
A:Title: Eliminated chromatin of Ascaris contains a gene that encodes a putative ribosome  
A:Reference number: A39106; MUID:91156655  
A:Accession: A39106  
A:Molecule type: mRNA  
A:Residues: 1-148 <ETT>  
A:Cross-references: GB:M59417; NID:q159650; PID:q159651  
C:Superfamily: rat ribosomal protein S19  
C:Keywords: protein biosynthesis; ribosome

Query Match 86.7%; Score 26; DB 2; Length 148;  
Best Local Similarity 83.3%; Pred. No. 48;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|||||  
DB 21 KFLKKS 26

## RESULT 14

S65460

apolipoprotein B - Atlantic salmon (fragment)  
C:Species: Salmo salar (Atlantic salmon)  
C:Date: 04-Dec-1997 #sequence\_revision 12-Dec-1997 #text\_change 12-Dec-1997  
C:Accession: S65460; I51362  
R:Babin, P.J.; Deryckere, F.; Gannon, F.  
Eur. J. Biochem. 230, 45-51, 1995  
A:Title: Presence of an extended duplication in the putative low-density-lipoprotein  
A:Reference number: I51362; MUID:95324549  
A:Accession: S65460  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1058 <BAB>  
A:Cross-references: EMBL:X81856; NID:q854619; PID:q854620  
A:Experimental source: liver  
C:Genetics:  
A:Gene: apob  
C:Superfamily: apolipoprotein B  
C:Keywords: calcium; cholesterol metabolism; chylomicron; glycoprotein; intestine; LD

Query Match 86.7%; Score 26; DB 2; Length 1058;  
Best Local Similarity 83.3%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|||||  
DB 1010 KFLKET 1015

## RESULT 15

S44283  
pail repressor homolog - Lactobacillus delbrueckii  
C:Species: Lactobacillus delbrueckii  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 04-Sep-1998  
C:Accession: S44283  
R:Klein, J.; Schmidt, U.; Plapp, R.  
submitted to the EMBL Data Library, October 1993  
A:Description: Cloning, heterologous expression, and sequencing of the proline iminop  
A:Reference number: S44282  
A:Accession: S44283  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-173 <KLE>  
A:Cross-references: EMBL:Z26948; NID:q506457; PID:q482922  
C:Superfamily: Bacillus subtilis transcription regulator

Query Match 86.7%; Score 26; DB 2; Length 173;  
Best Local Similarity 83.3%; Pred. No. 56;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|||||  
DB 40 KFLKET 45

Search completed: September 7, 1999, 23:22:47  
Job time: 954 sec

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	30	100.0	145	1	YPH2_MYCCA	P45614 mycoplasma
2	38	93.3	307	1	PE38_NVPQ	P45122 orgyia pseu
3	28	93.3	232	1	Y529_METJA	Q57949 methanococc
4	27	90.0	507	1	C7C3_ARATH	O64638 arabidopsis
5	27	90.0	412	1	CCA_SULSH	P77978 sulfolobus
6	27	90.0	410	1	COTR_CAVPO	P22323 cavia porce
7	27	90.0	1395	1	SP41_YEAST	P38904 saccharomyc
8	26	86.7	1094	1	DPOD_PLAFK	P30315 plasmodium
9	26	86.7	650	1	GY8B_MYCPN	P22447 mycoplasma
10	26	86.7	521	1	IC6A_HUMAN	Q92851 homo sapien
11	26	86.7	233	1	NCAP_BUNGE	P16993 bunyavirus
12	26	86.7	148	1	R19G_ASCSU	P24494 ascaris suu
13	26	86.7	173	1	YPIP_LACDL	P46543 lactobacill
14	26	86.7	659	1	YYBT_BACDU	P37484 bacillus su
15	25	83.3	713	1	AMVR_BACS8	P17692 bacillus sp
16	25	83.3	244	1	ARC3_CBCP	Q00901 clostridium
17	25	83.3	1355	1	ATC3_YEAST	P39524 saccharomyc
18	25	83.3	1006	1	BGAL_LACDE	P20043 lactobacill
19	25	83.3	825	1	BISC_HAEIN	P44798 haemophilus
20	25	83.3	815	1	CC53_YEAST	Q12018 saccharomyc
21	25	83.3	1033	1	CC68_KLUJA	Q00976 kluyveromyc
22	25	83.3	713	1	CDGT_BACSO	P05618 bacillus sp
23	25	83.3	712	1	CDGT_BAC33	P09121 bacillus sp
24	25	83.3	713	1	CDGT_BACSP	P30921 bacillus sp
25	25	83.3	713	1	CDGU_BACCI	P43379 bacillus ci
26	25	83.3	103	1	CHUB_EOUC	Q32214 equisetum s
27	25	83.3	391	1	CPSD_STRAG	Q04664 streptococc
28	25	83.3	569	1	CYSP_PLAFA	P25805 plasmodium
29	25	83.3	499	1	DLDH_YEAST	P09624 saccharomyc
30	25	83.3	1191	1	DNBI_MCMVS	P30672 murine cyto
31	25	83.3	1161	1	DP3A_BORBU	O51526 borrelia bu
32	25	83.3	430	1	EFIA_ENTHI	P31018 entamoeba h
33	25	83.3	521	1	GLGA_METJA	O59001 methanococc
34	25	83.3	795	1	HIS2_KLUJA	O13471 kluyveromyc
35	25	83.3	493	1	IMDH_STRPY	P50099 streptococc
36	25	83.3	1039	1	ITR4_MOUSE	Q00651 mus musculu
37	25	83.3	922	1	LCN2_NYCCO	P37609 lactococcus
38	25	83.3	1260	1	LINL_NYCCO	P08548 necticebus
39	25	83.3	600	1	MCM3_NAIZE	Q43704 zea mays (m
40	25	83.3	1774	1	MSAS_PENPA	P22367 penicillium
41	25	83.3	425	1	MUTL_AQUAE	O67518 aquifex aeo
42	25	83.3	351	1	NOV_CHTCK	P28686 gallus gall
43	25	83.3	353	1	NOV_COTJA	P42642 coturnix co

RT "Tandemly repeated sequence at the 3' end of the IE-2 gene of the  
RT baculovirus Orgyia pseudotsugata multicausid nuclear polyhedrosis  
RT virus is an enhancer element.";

```
RL VIROLOGY 187:97-106(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97271300.
RA AHRENS C.H., RUSSELL R.R., FUNK C.J., EVANS J., HARWOOD S.,
RA ROHMANN G.F.;
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RL VIROLOGY 229:381-399(1997).
CC -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M83827; G332542;
DR EMBL; U75930; G1911398;
DR PIR; A42191; WNV38.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
DR PFAM; PF00097; zf-C3HC4; 1.
KW EARLY PROTEIN; ZINC-FINGER; DNA-BINDING.
FT ZN_FING 39 91
SQ SEQUENCE 307 AA; 34714 MW; E712CE99 CRC32;

Query Match 93.3%; Score 28; DB 1; Length 307;
Best Local Similarity 83.3%; Pred. No. 22;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6
DB 110 KFIKKT 115

RESULT 3
Y529_METJA STANDARD; PRT; 232 AA.
AC Q57949;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0529.
GN MJ0529.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: WEAK, TO M.JANNASCHII MJ0173.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
```

```
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67502; G1591232;
DR TIGR; MJ0529;
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 232 AA; 27091 MW; 7E4228B8 CRC32;

Query Match 93.3%; Score 28; DB 1; Length 232;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6
DB 197 KFMKKT 202

RESULT 4
C7C3_ARATH STANDARD; PRT; 507 AA.
ID C7C3_ARATH
AC O64638;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CYTOCHROME P450 76C3 (EC 1.14.-.-).
GN CYP76C3 OR F17K2.11.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, COLUMBIA;
RA ROUNSLEY S.D., KAUL S., LIN X., KETCHUM K.A., CROSBY M.L.,
RA BRANDON R.C., SYKES S.M., MASON T.N., KERLAVAGE A.R., ADAMS M.D.,
RA SOMERVILLE C.R., VENTER J.C.;
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AC003680; G2979550;
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW OXIDOREDUCTASE; MONOOXYGENASE; HEME; MULTIGENE FAMILY.
FT BINDING 443 443 HEME (BY SIMILARITY).
SQ SEQUENCE 507 AA; 57285 MW; EE84B27 CRC32;

Query Match 90.0%; Score 27; DB 1; Length 507;
Best Local Similarity 83.3%; Pred. No. 61;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6
DB 123 RFLKKT 128

RESULT 5
CCA_SULSH STANDARD; PRT; 412 AA.
ID CCA_SULSH
AC F77978;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE TRNA NUCLEOTIDYLTRANSFERASE (EC 2.7.7.25) (TRNA ADENYLYLTRANSFERASE)
DE (TRNA CCA-PYROPHOSPHORYLASE) (CCA-ADDING ENZYME).
GN CCA.
```

OS SULFOLOBUS SHIBATAE.  
 OC ARCHAEA; CRENARCHAEOTA; SULFOLOBALES; SULFOLOBUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 96404875.  
 RA YUE D., MAIZELS N., WEINER A.M.:  
 RT "CCA-adding enzymes and poly(A) polymerases are all members of the  
 same nucleotidyltransferase superfamily: characterization of the  
 CCA-adding enzyme from the archaeal hyperthermophile Sulfolobus  
 shibatae.";  
 RL RNA 2:895-908(1996).  
 CC -1- FUNCTION: THIS ENZYME CARRIES OUT SYNTHESIS OF THE TRNA CCA  
 TERMINUS.  
 CC -1- CATALYTIC ACTIVITY: ATP + TRNA(N) - PYROPHOSPHATE + TRNA(N+1).  
 CC -1- SIMILARITY: BELONGS TO THE TRNA NUCLEOTIDYLTRANSFERASE / POLY(A)  
 POLYMERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U66004; G1575544; -.  
 KW TRANSFERASE; NUCLEOTIDYLTRANSFERASE; RNA-BINDING; TRNA PROCESSING.  
 SQ SEQUENCE 412 AA; 47872 MW; E8052BCA CRC32;  
 -----  
 Query Match 90.0%; Score 27; DB 1; Length 412;  
 Best Local Similarity 83.3%; Pred. No. 50;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KFLKKT 6  
 DB 402 KFLRKT 407  
 |||:|  
 -----  
 RESULT 6  
 COTR\_CAVPO STANDARD; PRT; 410 AA.  
 ID COTR\_CAVPO  
 AC P22323;  
 DT 01-AUG-1991 (REL. 19, CREATED)  
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
 DE CONTRAPSIN PRECURSOR (CP).  
 OS CAVIA PORCELLUS (GUINEA PIG).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; HYSTRICOGNATHI; CAVIIDAE; CAVIA.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-44.  
 RX MEDLINE: 91093294.  
 RA SUZUKI Y., YOSHIDA K., HONDA E., SINOHARA H.;  
 RT "Molecular cloning and sequence analysis of cDNAs coding for guinea  
 pig alpha 1-antitrypsinases S and F and contrapsin.";  
 RL J. BIOL. CHEM. 266:928-932(1991).  
 CC -1- FUNCTION: CONTRAPSIN INHIBITS TRYPSIN-LIKE PROTEASES.  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M57269; G191249; -.  
 DR PIR: C39088; C39088.  
 DR PROSITE: PS00284; SERPIN; 1.  
 DR PFAM: PF00079; serpin; 1.

DR HSSP: P01009; 8API.  
 KW SERPIN; SERINE PROTEASE INHIBITOR; PLASMA; GLYCOPROTEIN; SIGNAL.  
 FT SIGNAL 1 24  
 FT CHAIN 25 410 CONTRAPSIN.  
 FT ACT\_SITE 374 375 REACTIVE BOND (BY SIMILARITY).  
 FT CARBOHYD 62 62 POTENTIAL.  
 FT CARBOHYD 99 99 POTENTIAL.  
 FT CARBOHYD 162 162 POTENTIAL.  
 FT CARBOHYD 229 229 POTENTIAL.  
 FT CARBOHYD 263 263 POTENTIAL.  
 SQ SEQUENCE 410 AA; 45599 MW; EB64C2D7 CRC32;  
 -----  
 Query Match 90.0%; Score 27; DB 1; Length 410;  
 Best Local Similarity 83.3%; Pred. No. 50;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KFLKKT 6  
 DB 290 KFLRKT 295  
 |||:|  
 -----  
 RESULT 7  
 SP41\_YEAST STANDARD; PRT; 1395 AA.  
 ID SP41\_YEAST  
 AC P38904;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE SPP41 PROTEIN.  
 GN SPP41 OR YDR464W.  
 OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETES;  
 OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A364A;  
 RX MEDLINE: 94274035.  
 RA MADDOCK J.R., WEIDENHAMMER E.M., ADAMS C.C., LUNZ R.L.,  
 RA WOOLFORD J.L. JR.;  
 RT "Extragenic suppressors of sacccharomyces cerevisiae prp4 mutations  
 identify a negative regulator of PRP genes.";  
 RL GENETICS 136:833-847(1994).  
 CC -1- FUNCTION: NEGATIVE REGULATOR OF PRP3 AND PRP4 GENES. INTERACTS  
 CC WITH PRP8 AND RAP1.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U03673; G435023; -.  
 DR PIR: S47864; S47864.  
 DR SGD: L0002019; SPP41.  
 KW NUCLEAR PROTEIN.  
 FT DOMAIN 683 699 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 SQ SEQUENCE 1395 AA; 156933 MW; 4C8A22E9 CRC32;  
 -----  
 Query Match 90.0%; Score 27; DB 1; Length 1395;  
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KFLKKT 6  
 DB 1216 KYLAKT 1221  
 |||:|  
 -----  
 RESULT 8

DPDOD\_PLAFK  
 ID DPDOD\_PLAFK STANDARD; PRT: 1094 AA.  
 AC P30315;  
 DT 01-APR-1993 (REL. 25, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE DNA POLYMERASE DELTA CATALYTIC CHAIN (EC 2.7.7.7).  
 GN POLD.  
 OS PLASMODIUM FALCIPARUM (ISOLATE K1 / THAILAND).  
 OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDIA; PLASMODIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92107655.  
 RA RIDLEY R.G., WHITE J.H., MCALLESE S.M., GOMAN M., ALANO P.,  
 RA DEVRIES E., KILBEY B.J.;  
 RT "DNA polymerase delta: gene sequences from Plasmodium falciparum  
 RT indicate that this enzyme is more highly conserved than DNA  
 RT polymerase alpha.";  
 RL NUCLEIC ACIDS RES. 19:6731-6736(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92131065.  
 RA FOX B.A., BZIK D.J.;  
 RT "The primary structure of Plasmodium falciparum DNA polymerase delta  
 RT is similar to drug sensitive delta-like viral DNA polymerases.";  
 RL MOL. BIOCHEM. PARASITOL. 49:289-296(1991).  
 CC -!- FUNCTION: THIS POLYMERASE POSSESS TWO ENZYMAIC ACTIVITIES:  
 CC DNA SYNTHESIS (POLYMERASE) AND AN EXONUCLEOTIC ACTIVITY THAT  
 CC DEGRADES SINGLE STRANDED DNA IN THE 3' TO 5' DIRECTION.  
 CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -  
 CC N PYROPHOSPHATE + DNA(N).  
 CC -!- SUBUNIT: HETERODIMER WITH SUBUNITS OF 125 KD AND 50 KD. THE 125 KD  
 CC SUBUNIT CONTAINS THE POLYMERASE ACTIVE SITE AND MOST LIKELY THE  
 CC ACTIVE SITE FOR THE 3'-5' EXONUCLEASE ACTIVITY.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES: ALPHA, BETA, GAMMA,  
 CC DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR DIFFERENT REACTIONS  
 CC OF DNA SYNTHESIS.  
 CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M64715; G160270; -  
 DR EMBL; X62423; G9941; -  
 DR EMBL; M63941; G160268; -  
 DR PIR; A45639; A45639.  
 DR PROSITE; PS00116; DNA\_POLYMERASE\_B; 1.  
 DR PFAM; PF00136; DNA\_POL\_B; 1.  
 DR TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION;  
 KW DNA-BINDING; HYDROLASE; EXONUCLEASE; ZINC-FINGER; NUCLEAR PROTEIN.  
 FT ZN-FING 1003 1019 C4-TYPE (POTENTIAL).  
 FT ZN-FING 1049 1067 C4-TYPE (POTENTIAL).  
 SQ SEQUENCE 1094 AA; 126883 MW; 4EF5EE3E CRC32;

Query Match 86.7%; Score 26; DB 1; Length 1094;  
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 Db 995 KFLKKS 1000  
 |||||

RESULT 9  
 GYRB\_MYCPN  
 ID GYRB\_MYCPN STANDARD; PRT: 650 AA.

P22447;  
 AC 01-AUG-1991 (REL. 19, CREATED)  
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE DNA GYRASE SUBUNIT B (EC 5.99.1.3).  
 GN GYRB.  
 OS MYCOPLASMA PNEUMONIAE.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;  
 OC MYCOPLASMATACEAE; MYCOPLASMA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91041721.  
 RA COLMAN S.D., HU P.C., BOTT K.F.;  
 RT "Mycoplasma pneumoniae DNA gyrase genes.";  
 RL MOL. MICROBIOL. 4:1129-1134(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE; 97105885.  
 RA HIMMELREICH R., HILBERT H., PIRKL E., LI B.-C.,  
 RA HERRMANN R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).  
 CC -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-  
 CC STRANDED DNA IN AN ATP-DEPENDANT MANNER AND ALSO CATALYZES THE  
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED  
 CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.  
 CC -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING  
 CC OF DOUBLE-STRANDED DNA.  
 CC -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA  
 CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE  
 CC ENZYME FORMS AN A2B2 TETRAMER.  
 CC -!- SIMILARITY: BELONGS TO TYPE II TOPOISOMERASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X53555; G44484; -  
 DR EMBL; AE000016; G1673811; -  
 DR PIR; S11767; ISYMEP  
 DR PROSITE; PS00177; TOPOISOMERASE\_II; 1.  
 DR PFAM; PF00204; DNA\_topoisoII; 1.  
 DR PFAM; PF00986; DNA\_gyraseB\_C; 1.  
 DR HSP; P06982; LAJ6.  
 DR TOPOISOMERASE; ISOMERASE; ATP-BINDING; ANTIBIOTIC RESISTANCE.  
 KW TOPOISOMERASE; ISOMERASE; ATP-BINDING; ANTIBIOTIC RESISTANCE.  
 SQ SEQUENCE 650 AA; 73809 MW; D7CE7F9B CRC32;

Query Match 86.7%; Score 26; DB 1; Length 650;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 Db 318 KFLKET 323  
 |||||

RESULT 10  
 ICEA\_HUMAN  
 ID ICEA\_HUMAN STANDARD; PRT: 521 AA.

AC Q92851; Q99845;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE CASPASE-10 PRECURSOR (EC 3.4.22.-) (ICE-LIKE APOPTOTIC PROTEASE 4)  
 DE (APOPTOTIC PROTEASE MCH-4) (FAS-ASSOCIATED DEATH DOMAIN PROTEIN  
 DE INTERLEUKIN-1B-CONVERTING ENZYME 2) (FLICE2).

GN CASP10 OR MCH4.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A. (SHORT FORM).  
 RC TISSUE-T CELL;  
 RX MEDLINE; 9653838.  
 RA FERNANDES-ALNEMRI T., ARMSTRONG R.C., KREBS J., SRINIVASULA S.M.,  
 RA WANG L., BULLRICH F., FRITZ L.C., TRAPANI J.A., TOMASELLI K.J.,  
 RA LITWACK G., ALNEMRI E.S.;  
 RA "In vitro activation of CPP32 and Mch3 by Mch4, a novel human  
 RT apoptotic cysteine protease containing two FADD-like domains";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 93:7464-7469(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A. (LONG FORM).  
 RX MEDLINE; 97197836.  
 RA VINCENZ C., DIXIT V.M.;  
 RA "Fas-associated death domain protein interleukin-lb $\beta$ -converting in  
 RT enzyme 2 (FLICE2), an ICE/Ced-3 homologue, is proximally involved in  
 RT CD95- and p55-mediated death signaling.";  
 RL J. BIOL. CHEM. 272:6578-6583(1997).  
 RN [3]  
 RP PARTIAL SEQUENCE, AND PROCESSING.  
 RX MEDLINE; 97121412.  
 RA SRINIVASULA S.M., AHMAD M., FERNANDES-ALNEMRI T., LITWACK G.,  
 RA ALNEMRI E.S.;  
 RA "Molecular ordering of the Fas-apoptotic pathway: the Fas/APO-1  
 RT protease Mch5 is a CrmA-inhibitable protease that activates multiple  
 RT Ced-3/ICE-like cysteine proteases";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 93:14486-14491(1996).  
 CC -1- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES  
 CC RESPONSIBLE FOR APOPTOSIS EXECUTION. RECRUITED TO BOTH FAS- AND  
 CC TNFR-1 RECEPTORS IN A FADD DEPENDENT MANNER. MAY PARTICIPATE IN  
 CC THE GRANZYME B APOPTOTIC PATHWAYS. CLEAVES AND ACTIVATES CASPASE-  
 CC 3, -4, -6, -7, -8, AND -9. HYDROLYZES THE SMALL- MOLECULE  
 CC SUBSTRATES, TYR-VAL-ALA-ASP-|-AMC AND ASP-GLU-VAL-ASP-|-AMC.  
 CC -1- SUBUNIT: HETERODIMER OF A 23/17 KD (P23/17) DEPENDING ON THE  
 CC SPLICING EVENTS AND A 12 KD (P12) SUBUNIT.  
 CC -1- TISSUE SPECIFICITY: DETECTABLE IN MOST TISSUES. LOWEST EXPRESSION  
 CC IS SEEN IN BRAIN, KIDNEY, PROSTATE, TESTIS, AND COLON.  
 CC -1- DOMAIN: CONTAINS TWO DEATH EFFECTOR DOMAINS (ALSO CALLED MORT  
 CC MODULES), WHICH ARE INVOLVED IN BINDING TO THE ADAPTOR MOLECULE  
 CC FADD.  
 CC -1- PTM: CLEAVAGE BY GRANZYME B AND AUTOCATALYTIC ACTIVITY GENERATE  
 CC THE TWO ACTIVE SUBUNITS.  
 CC -1- ALTERNATIVE PRODUCTS: A LONG FORM (SHOWN HERE) AND A SHORT FORM  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14; ALSO KNOWN AS THE  
 CC CASPASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U60519; G1498324; -.  
 CC EMBL; U86214; G1835779; -.  
 CC MIM; 601762; -.  
 CC PROSITE; PS01121; CASPASE\_HIS; 1.  
 CC PROSITE; PS01122; CASPASE\_CYS; 1.  
 CC PFAM; PF00655; ICE\_p10; 1.  
 CC PFAM; PF00656; ICE\_p20; 1.  
 CC PFAM; PF01335; CED; 2.  
 CC HSP; P42574; IPAU.  
 CC HYDROLASE; THIOL PROTEASE; APOPTOSIS; ZYMOGEN; ALTERNATIVE SPLICING.  
 FT PROPEP 1 219 CASPASE-10 SUBUNIT P23/17.  
 FT CHAIN 220 415 CASPASE-10 SUBUNIT P12.  
 FT CHAIN 416 521 CASPASE-10 SUBUNIT P12.  
 FT DOMAIN 20 96 DEATH EFFECTOR DOMAIN.

FT DOMAIN 115 187 DEATH EFFECTOR DOMAIN.  
 FT ACT\_SITE 358 358 BY SIMILARITY.  
 FT ACT\_SITE 401 401 BY SIMILARITY.  
 FT VARSPPLIC 229 271 MISSING (IN SHORT FORM).  
 FT VARSPPLIC 473 521 MKLEKTEMEIRGRKRTVMGAKQISATSLPTAISAOPTPRP  
 FT PMRRSSVS -> HEDILSILTAVNDYVRRVYDKGKTQKQ  
 FT MPQAPATLRKKLVFPVPLDALSI (IN SHORT FORM).  
 FT CONFLICT 68 68  
 SQ SEQUENCE 521 AA; 58878 MW; D4D64D28 CRC32;  
 Query Match 86.7%; Score 26; DB 1; Length 521;  
 Best Local Similarity 83.3%; Pred. No. 1;e-02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KFLKKT 6  
 DB 475 KFLKKT 480  
 III:II  
 RESULT 11  
 NCAP\_BUNGE STANDARD; PRT; 233 AA.  
 ID NCAP\_BUNGE STANDARD; PRT; 233 AA.  
 AC P16993;  
 DT 01-AUG-1990 (REL. 15, CREATED)  
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1990 (REL. 15, LAST ANNOTATION UPDATE)  
 DE NUCLEOCAPSID PROTEIN (NUCLEOPROTEIN).  
 GN N.  
 OS BUNYAVIRUS GERMISTON.  
 OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; BUNYAVIRIDAE; BUNYAVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88019678.  
 RA GERBAUD S., VIALAT P., PARDIGON N., WYCHOWSKI C., GIRARD M.,  
 RA BOULOY M.;  
 RA "The S segment of the Germiston virus RNA genome can code for three  
 RT proteins";  
 RL VIRUS RES. 8:1-13(1987).  
 CC -1- SUBCELLULAR LOCATION: INTERNAL PROTEIN OF VIRUS PARTICLE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M19420; G210758; -.  
 CC PFAM; PF00952; Bunya\_nucleocap; 1.  
 KW NUCLEOCAPSID; RNA-BINDING.  
 SQ SEQUENCE 233 AA; 26703 MW; B4393642 CRC32;  
 Query Match 86.7%; Score 26; DB 1; Length 233;  
 Best Local Similarity 83.3%; Pred. No. 48;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KFLKKT 6  
 DB 175 KFLKKT 180  
 III:II  
 RESULT 12  
 R19G\_ASCSU STANDARD; PRT; 148 AA.  
 ID R19G\_ASCSU STANDARD; PRT; 148 AA.  
 AC P24494;  
 DT 01-MAR-1992 (REL. 21, CREATED)  
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1992 (REL. 33, LAST ANNOTATION UPDATE)  
 DE 40S RIBOSOMAL PROTEIN S19G (ELIMINATED PROTEIN NO. 1).  
 GN RPS19G OR ALEP-1.

```

OS ASCARIS SUUM (PIG ROUNDWORM) (ASCARIS LUMBRICOIDES).
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; ASCARIDIDA;
OC ASCARIDOIDEA; ASCARIDIDAE; ASCARIS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91156655.
RA ETTER A., ABOUTANOS M., TOBLER H., MUELLER F.;
RT "Eliminated chromatin of Ascaris contains a gene that encodes a
RT putative ribosomal protein."
RL PROC. NATL. ACAD. SCI. U.S.A. 88:1593-1596(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA ETTER A.;
RL SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: ELIMINATION OF THE ALEP-1 GENE FROM ALL SOMATIC CELLS IN
CC ITS FULLY ACTIVE STATE MAY REPRESENT AN ALTERNATIVE WAY TO GENE
CC REGULATION.
CC -!- DEVELOPMENTAL STAGE: OOCYTES AND EARLY EMBRYONIC STAGES UP TO ITS
CC ELIMINATION.
CC -!- SIMILARITY: BELONGS TO THE S19E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59417; G159651; -
DR EMBL; X75543; G414609; -
DR PIR; A39106; A39106.
DR PROSITE; PS00628; RIBOSOMAL_S19E; 1.
DR PFAM; PF01090; S19e; 1.
KW RIBOSOMAL PROTEIN; MULTIGENE FAMILY.
FT VARIANT 33 E->D.
SQ SEQUENCE 148 AA; 16584 MW; 32722282 CRC32;
-----
Query Match 86.7%; Score 26; DB 1; Length 148;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6
Db 21 KFLKKS 26
-----
RESULT 13
ID YPIP_LACDL STANDARD; PRT; 173 AA.
AC P46543;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 19.8 KD PROTEIN IN PEPI 3 REGION.
OS LACTOBACILLUS DELBRUECKII (SUBSP. LACTIS).
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; LACTOBACILLACEAE;
OC LACTOBACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-DSM 7290.
RX MEDLINE; 94297735.
RA KLEIN J.R., SCHMIDT U., PLAPP R.;
RT "Cloning, heterologous expression, and sequencing of a novel proline
RT iminopeptidase gene, pepI, from Lactobacillus delbrueckii subsp.
RT lactis DSM 7290."
RL MICROBIOLOGY 140:1133-1139(1994).
CC -!- SIMILARITY: STRONG, TO B.SUBTILIS PAIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z26948; G482922; -
DR PFAM; PF00583; Acetyltransf; 1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 173 AA; 19819 MW; 569164BD CRC32;
-----
Query Match 86.7%; Score 26; DB 1; Length 173;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6
Db 40 KFLKKT 45
-----
RESULT 14
ID YBET_BACSU STANDARD; PRT; 659 AA.
AC P37484;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 74.3 KD PROTEIN IN RPLI-COTF INTERGENIC REGION.
GN YBET.
OS BACILLUS SUBTILIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC BACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-168;
RX MEDLINE; 96051385.
RA OGASAWARA N., NAKAI S., YOSHIKAWA H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin."
RL DNA RES. 1:1-14(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D26185; G467336; -
DR EMBL; Z99124; E1184777; -
DR SUBTILIST; BG10010; YBET.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 659 AA; 74325 MW; 2CE6AD2E CRC32;
-----
Query Match 86.7%; Score 26; DB 1; Length 659;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6
Db 530 KFLKET 535
-----
RESULT 15
ID AMYR_BACS8 STANDARD; PRT; 713 AA.
AC P17692;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE RAW-STARCH-DIGESTING AMYLASE PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-

```



```

DE GLUCAN GLUCANOHYDROLASE).
OS BACILLUS SP. (STRAIN B1018).
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC BACILLUS.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-47.
RX MEDLINE; 90147765.
RA IYKOR P., TSUKAGOSHI N., UDAKA S.;
RT "Nucleotide sequence of the raw-starch-digesting amylase gene from
RT Bacillus sp. B1018 and its strong homology to the cyclodextrin
RT glucanotransferase genes.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 166:630-636(1990).
CC -1- FUNCTION: THIS ENDO-TYPE ADSORBABLE AMYLASE IS CAPABLE TO
CC DIGEST RAW-STARCH.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33302; G142509; -
DR EMBL; D90112; G216325; -
DR PIR; A34648; A34648.
DR PIR; S09196; S09196.
DR PFAM; PF00128; alpha-amylase; 1.
DR PFAM; PF00686; CBD_2; 1.
DR HSP; P43379; ICXF.
DR KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; CALCIUM; SIGNAL.
FT SIGNAL 1 27
FT CHAIN 28 713 RAW-STARCH DIGESTING AMYLASE.
FT ACT_SITE 256 256 BY SIMILARITY.
FT ACT_SITE 260 260 BY SIMILARITY.
FT ACT_SITE 355 355 BY SIMILARITY.
SQ SEQUENCE 713 AA; 77420 MW; 5C83248D CRC32;

```

```

Query Match      83.3%; Score 25; DB 1; Length 713;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 KFLKK 5
    |||||
Db 678 KFLKK 682

```

```

Search completed: September 7, 1999, 23:59:13
Job time: 525 sec

```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:42 ; Search time 160.2 Seconds  
(without alignments)  
2.305 Million cell updates/sec

Title: US-09-124-280A-32

Perfect score: 30

Sequence: 1 KFLKKT 6

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10.\*

- 1: sp-archaea.\*
- 2: sp-bacteria.\*
- 3: sp-fungi.\*
- 4: sp-human.\*
- 5: sp-invertebrate.\*
- 6: sp-mammal.\*
- 7: sp-mhc.\*
- 8: sp-organelle.\*
- 9: sp-phase.\*
- 10: sp-plant.\*
- 11: sp-rodent.\*
- 12: sp-virus.\*
- 13: sp-vertebrate.\*
- 14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	30	100.0	764	10	O64607	O64607 arabidopsis
2	28	93.3	537	3	O59831	O59831 schizosach
3	28	93.3	743	4	O43373	O43373 homo sapien
4	28	93.3	622	5	Q21389	Q21389 caenorhabdi
5	28	93.3	609	10	Q92QA8	Q92QA8 arabidopsis
6	27	90.0	433	2	P71096	P71096 bacteroides
7	27	90.0	293	2	O50644	O50644 lactococcus
8	27	90.0	503	2	O30417	O30417 lactococcus
9	27	90.0	202	2	O84499	O84499 chlamydia t
10	27	90.0	1435	3	Q03291	Q03291 saccharomyc
11	27	90.0	761	5	O16947	O16947 caenorhabdi
12	27	90.0	211	5	O17556	O17556 caenorhabdi
13	27	90.0	213	5	Q23131	Q23131 caenorhabdi
14	27	90.0	189	10	O49581	O49581 arabidopsis
15	27	90.0	993	10	O82317	O82317 arabidopsis
16	27	90.0	178	10	Q92Q28	Q92Q28 arabidopsis
17	27	90.0	178	12	O87505	O87505 chimpanzee
18	27	90.0	422	12	Q9YVW4	Q9YVW4 melanoplus
19	26	86.7	396	2	O30714	O30714 helicobacte
20	26	86.7	813	2	O31558	O31558 borrelia bu
21	26	86.7	250	2	O67647	O67647 aquifex aeo
22	26	86.7	192	2	O44599	O44599 buchnera ap
23	26	86.7	457	2	O48902	O48902 mycoplasma
24	26	86.7	305	2	Q92CE0	Q92CE0 rickettsia
25	26	86.7	312	2	Q92729	Q92729 chlamydia p
26	26	86.7	1447	3	Q09140	Q09140 schizosach
27	26	86.7	327	5	Q94244	Q94244 caenorhabdi
28	26	86.7	1827	5	Q20535	Q20535 caenorhabdi
29	26	86.7	3396	5	O62263	O62263 caenorhabdi

#### ALIGNMENTS

##### RESULT 1

O64607  
ID O64607 PRELIMINARY; PRT; 764 AA.  
AC O64607;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
DE F1707.17 PROTEIN.  
GN F1707.17.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA VYSOTSKAIA V.S., SCHWARTZ J.R., TORIUMI M., YU G., KWAN A., QJI O.,  
RA LIU S., LI J., ARAUJO R., AU M., BRENDEN V., BUEHLER E., CONWAY A.B.,  
RA CONWAY A.R., DEWAR K., FENG J., KIM C., KURTZ D., LI Y., PALM C.J.,  
RA SHINN P., SUN H., DAVIS R.W., ECKER J.R., FEDERSPIEL N.A.,  
RA THEOLOGIS A.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

##### RESULT 2

O64607  
ID O64607 PRELIMINARY; PRT; 764 AA.  
AC O64607;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
DE F1707.17 PROTEIN.  
GN F1707.17.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA VYSOTSKAIA V.S., SCHWARTZ J.R., TORIUMI M., YU G., KWAN A., QJI O.,  
RA LIU S., LI J., ARAUJO R., AU M., BRENDEN V., BUEHLER E., CONWAY A.B.,  
RA CONWAY A.R., DEWAR K., FENG J., KIM C., KURTZ D., LI Y., PALM C.J.,  
RA SHINN P., SUN H., DAVIS R.W., ECKER J.R., FEDERSPIEL N.A.,  
RA THEOLOGIS A.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA THEOLOGIS A.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA THEOLOGIS A.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA THEOLOGIS;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC003671; AAC18809.1; -.  
SQ SEQUENCE 764 AA; 85199 MW; 8D4C0AD2 CRC32;

Query Match 100.0%; Score 30; DB 10; Length 764;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKKT 6

DB 243 KFLKKT 248

OS9831  
ID O59831 PRELIMINARY; PRT; 537 AA.  
AC O59831;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE AMINO ACID PERMEASE.  
GN SPCC965.11C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H-;  
RA WOOD V., RAJADREAM M.A., BARRELL B.G., RIEGER M;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL023590; CAA19071.1; -;  
DR PFAM; PF00324; aa-permeases; 1.  
SQ SEQUENCE 537 AA; 59431 MW; 42F219F7 CRC32;

Query Match 93.3%; Score 28; DB 3; Length 537;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
DB 485 KFIKKT 490  
II:III

RESULT 3  
ID O43373 PRELIMINARY; PRT; 743 AA.  
AC O43373;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE PAC CLONE DJ0897G10 FROM 7Q11.23-Q21.1, COMPLETE SEQUENCE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KALICKI J., ELLIOTT G.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA WATERSTON R.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC004082; AAB97937.1; -;  
DR PFAM; PF00168; C2; 2.  
SQ SEQUENCE 743 AA; 81865 MW; 4A4DB2BF CRC32;

Query Match 93.3%; Score 28; DB 4; Length 743;  
Best Local Similarity 83.3%; Pred. No. 2.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
DB 706 KFMKKT 711  
II:III

RESULT 4  
ID Q21389 PRELIMINARY; PRT; 622 AA.  
AC Q21389;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
DE K09C8.4 PROTEIN.  
GN K09C8.4.

OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KERSHAW J.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WAYSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans";  
RL Nature 368:32-38(1994).  
DR EMBL; 268006; CAA91997.1; -;  
SQ SEQUENCE 622 AA; 73475 MW; 269A0E31 CRC32;

Query Match 93.3%; Score 28; DB 5; Length 622;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
DB 608 KFMKKT 613  
II:III

RESULT 5  
ID Q9ZQA8 PRELIMINARY; PRT; 609 AA.  
AC Q9ZQA8;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE PUTATIVE POLYA BINDING PROTEIN.  
GN F13K3.6.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA LIN X., KAUL S., SHEA T.P., FUJII C.Y., SHEN M., VANAKEN S.E.,  
RA BARNSTEAD M.E., MASON T.M., BOWMAN C.L., RONNING C.M., BENITO M.,  
RA CARRERA A.J., CREASY T.H., BUELL C.R., TOWN C.D., NIERMAN W.C.,  
RA FRASER C.M., VENTER J.C.;  
RT "Arabidopsis thaliana chromosome II BAC F13K3 genomic sequence";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006282; AAD20142.1; -;  
SQ SEQUENCE 609 AA; 67917 MW; 44683B6D CRC32;

Query Match 93.3%; Score 28; DB 10; Length 609;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
DB 186 KFMKKT 191  
II:III

RESULT 6

P71096  
ID P71096 PRELIMINARY; PRT; 433 AA.  
AC P71096;  
DT 01-FEB-1997 (TReMBLrel. 02, Created)  
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)  
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
DE REGULATORY PROTEIN.  
GN SUSA.  
OS Bacteroides thetaiotaomicron.  
OC Bacteria; Cytophagales; Bacteroidaceae; Bacteroides.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5842;  
RA D'ELIA J., SALVERS A.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U67062; AAB39215.1; -;  
SQ SEQUENCE 433 AA; 50646 MW; 13D3D4D5 CRC32;

Query Match 90.0%; Score 27; DB 2; Length 433;  
Best Local Similarity 83.3%; Pred. No. 2.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|:||||  
Db 76 KYLAKT 81

RESULT 7  
O50644 PRELIMINARY; PRT; 293 AA.  
ID O50644  
AC O50644;  
DT 01-JUN-1998 (TReMBLrel. 06, Created)  
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
DE GAD C (FRAGMENT).  
OS Lactococcus lactis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=01-7;  
RA NOMURA M., FUJITA Y., KOBAYASHI M., OOMOMO S.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB010789; BAA24584.1; -;  
FT NON\_TER 1  
SQ SEQUENCE 293 AA; 32569 MW; D8E8DA5B CRC32;

Query Match 90.0%; Score 27; DB 2; Length 293;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|:||||  
Db 110 KFLRKT 115

RESULT 8  
O30417 PRELIMINARY; PRT; 503 AA.  
ID O30417  
AC O30417;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE GADC.  
OS Lactococcus lactis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
RN [1]  
RP SEQUENCE OF 1-119 FROM N.A.  
RC STRAIN=MGI1363;

RX MEDLINE; 98265928.  
RA SANDERS J.W., VENEMA G., KOK J., LEENHOUTS K.;  
RT "Identification of a sodium chloride-regulated promoter in  
RT Lactococcus lactis by single-copy chromosomal fusion with a reporter  
RT gene";  
RL Mol. Gen. Genet. 257:681-685(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MGI1363;  
RX MEDLINE; 98143417.  
RA SANDERS J.W., LEENHOUTS K., BURGHOOHN J., BRANDS J.R., VENEMA G.,  
RA KOK J.;  
RT "A chloride-inducible acid resistance mechanism in Lactococcus lactis  
RT and its regulation.";  
RL Mol. Microbiol. 27:299-310(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MGI1363;  
RA SANDERS J.W., VENEMA G., KOK J., LEENHOUTS K.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF005098; AAC46187.1; -;  
DR PFAM: PF00324; aa\_permeases; 1.  
SQ SEQUENCE 503 AA; 55368 MW; 479F8FEF CRC32;

Query Match 90.0%; Score 27; DB 2; Length 503;  
Best Local Similarity 83.3%; Pred. No. 2.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|:||||  
Db 320 KFLRKT 325

RESULT 9  
O84499 PRELIMINARY; PRT; 202 AA.  
ID O84499  
AC O84499;  
DT 01-NOV-1998 (TReMBLrel. 08, Created)  
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
DE PREDICTED PHOSPHATASE/KINASE.  
GN YACE.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D/UW-3/CX;  
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,  
RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,  
RA DAVIS R.W.;  
RT "Genome Sequence of an Obligate Intracellular Pathogen of Humans:  
RT Chlamydia trachomatis";  
RL Science 0:0-0(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D/UW-3/CX;  
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,  
RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,  
RA DAVIS R.W.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE001322; AAC68092.1; -;  
DR PFAM: PF01121; UPF0038; 1.  
SQ SEQUENCE 202 AA; 22953 MW; 3B43E07F CRC32;

Query Match 90.0%; Score 27; DB 2; Length 202;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
|:||||  
Db 145 RFLKKT 150

RESULT 10  
Q03291  
ID Q03291 PRELIMINARY; PRT: 1435 AA.  
AC Q03291;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
DE REGULATORY PROTEIN.  
GN SPP41  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA DIETRICH F.S., MULLIGAN J., ALLEN E., ARAUJO R., AVILES E., BERNO A.,  
RA CARPENTER J., CHEN E., CHERRY J.M., CHUNG E., DUNCAN M.,  
RA HUNICKE-SMITH S., HYMAN R., KOMP C., LASHKARI D., LEW H., LIN D.,  
RA MOSEDALE D., NAKAHARA K., NAMATH A., OEFNER P., OH C., PETEL F.X.,  
RA ROBERTS D., SCHRAMM S., SCHROEDER M., SHOGREN T., SHROFF N.,  
RA WINANT A., YELTON M., BOTSTEIN D., DAVIS R.W.;  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA DIETRICH F.S.;  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA JIA Y., CHERRY J.M.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U33050; AAB64901.1; -;  
DR SGD; L0002019; SPP41.  
SQ SEQUENCE 1435 AA; 161596 MW; 9341784B CRC32;

Query Match 90.0%; Score 27; DB 3; Length 1435;  
Best Local Similarity 83.3%; Pred. No. 6.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
I:||||  
DB 1216 KYLKKT 1221

RESULT 11  
O16947  
ID O16947 PRELIMINARY; PRT: 761 AA.  
AC O16947;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE K09C6.1 PROTEIN.  
GN K09C6.1  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";  
RN [2]  
RP SEQUENCE FROM N.A.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";

RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA GOELA D., HARPER M.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA WATERSTON R.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF022975; AAB70668.1; -;  
SQ SEQUENCE 761 AA; 87480 MW; DDC23304 CRC32;

Query Match 90.0%; Score 27; DB 5; Length 761;  
Best Local Similarity 83.3%; Pred. No. 3.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
I:||||  
DB 30 RFLAKT 35

RESULT 12  
O17556  
ID O17556 PRELIMINARY; PRT: 211 AA.  
AC O17556;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-JAN-1999 (TReMBLrel. 09, Last annotation update)  
DE C01B9.2 PROTEIN.  
GN C01B9.2  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA PERCY C.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";  
RN [3]  
RP SEQUENCE FROM N.A.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";

Query Match 90.0%; Score 27; DB 5; Length 211;  
Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
I:||||  
DB 6 KYLKKT 11

RESULT 13  
Q23131  
ID Q23131 PRELIMINARY; PRT: 213 AA.  
AC Q23131;

DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TReMBLrel. 09, Last annotation update)  
 DE W02H3.2 PROTEIN.  
 GN W02H3.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA COTTAGE A.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL; 268013; CAA92021.1; -;  
 SQ SEQUENCE 213 AA; 25547 MW; 3FC6C424 CRC32;

Query Match 90.0%; Score 27; DB 5; Length 213;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 DB 166 KFLKRT 171

RESULT 14  
 ID 049581 PRELIMINARY; PRT; 189 AA.  
 AC 049581;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
 DE AUXIN INDUCED LIKE-PROTEIN.  
 GN F8F16.140.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eukaryotes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA BEVAN M., BRANDT P., DOSE S., JARKE D., SCHAREE M., SCHON O.,  
 RA HOHEISEL J., MEWES H.W., MAYER K., SCHUELLER C.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RA EU ARABIDOPSIS SEQUENCING PROJECT;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL021633; CAA16535.1; -;  
 SQ SEQUENCE 189 AA; 21417 MW; 83BD79C8 CRC32;

Query Match 90.0%; Score 27; DB 10; Length 189;  
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6

DB 66 KFLKRT 71  
 IIII:|

RESULT 15  
 ID 082317 PRELIMINARY; PRT; 993 AA.  
 AC 082317;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE F17H15.17 PROTEIN.  
 GN F17H15.17.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eukaryotes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA STRAIN-CL. COLUMBIA;  
 RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,  
 RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;  
 RT "Arabidopsis thaliana chromosome II BAC F17H15 genomic sequence";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC005395; AAC42250.1; -;  
 SQ SEQUENCE 993 AA; 110549 MW; DFB51C2C CRC32;

Query Match 90.0%; Score 27; DB 10; Length 993;  
 Best Local Similarity 83.3%; Pred. No. 4.7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
 DB 981 RFLKKT 986

Search completed: September 7, 1999, 22:47:44  
 Job time: 7964 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:04 ; Search time 135.78 Seconds  
(without alignments)  
1.744 Million cell updates/sec

Title: US-09-124-280A-33

Perfect score: 58

Sequence: 1 CKKLFCKTK 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	58	100.0	10	1	R33529
2	58	100.0	10	1	R39290
3	58	100.0	10	1	R21621
4	53	91.4	9	1	R33530
5	53	91.4	9	1	R39291
6	53	91.4	9	1	R21622
7	35	60.3	2783	1	R23962
8	35	60.3	2783	1	R23963
9	35	60.3	1088	1	R19786
10	34	58.6	333	1	R46299
11	33	56.9	551	1	R06646
12	33	56.9	551	1	R07506
13	32	55.2	1706	1	R92103
14	32	55.2	1141	1	R44777
15	32	55.2	10	1	R71789
16	31.5	54.3	10	1	R21606
17	31.5	54.3	10	1	R25270
18	31	53.4	650	1	R39346
19	31	53.4	122	1	R77193
20	31	53.4	475	1	R77194
21	31	53.4	475	1	R95595
22	31	53.4	127	1	R38756
23	31	53.4	1726	1	R29658
24	31	53.4	178	1	R81513
25	31	53.4	554	1	R81514
26	31	53.4	541	1	R89954
27	31	53.4	232	1	R83318
28	31	53.4	944	1	R43662
29	30.5	52.6	3396	1	R43662
30	30.5	52.6	798	1	R06592
31	30.5	52.6	37	1	R33115
32	30	51.7	32	1	R15252
33	30	51.7	514	1	R15237
34	30	51.7	450	1	R30738
35	30	51.7	336	1	R30740
36	30	51.7	336	1	R34293
37	30	51.7	432	1	R34294
38	30	51.7	739	1	R41334
39	30	51.7	701	1	R41335
40	30	51.7	434	1	R57434
41	30	51.7	750	1	R72078
42	30	51.7	712	1	R72079
43	30	51.7	749	1	R72080

ALIGNMENTS

RESULT 1

R33529 R33529 standard; peptide; 10 AA.  
AC R33529;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..7  
PN 2A9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992: 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 6; Page 32; 39pp; English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 1; Length 10;

Best Local Similarity 100.0%; Pred No. 0.00079; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10

Db 1 CKKLFCKTK 10

RESULT 2

R39290 R39290 standard; peptide; 10 AA.

ID R39290;  
AC R39290;  
DT 22-DEC-1993 (first entry)  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
KW quantification.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..7  
PN W09314115-A.

PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 6; Page 32; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00079;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10  
 |||||  
 Db 1 CKKLFCKTK 10

## RESULT 3

ID W21621 standard; peptide; 10 AA.  
 AC W21621;  
 DT 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #33.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.

FH Key Location/Qualifiers

FT disulfide bond 1. .7  
 PN W09638163-A1.  
 PD 05-DEC-1996.

PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;  
 DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 37; Page 28; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00079;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10  
 |||||  
 Db 1 CKKLFCKTK 10

## RESULT 4

ID R33530 standard; peptide; 9 AA.  
 AC R33530;  
 DT 07-JUL-1993 (first entry)

DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.

FH Key Location/Qualifiers

FT disulfide bond 1. .7  
 PN ZA9200943-A.

PD 25-NOV-1992.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORRO M.

PI Porro M;  
 DR WPI; 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 7; Page 32; 39pp; English.

CC This peptide is a specific example of a generic peptide of

CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 9 AA;

Query Match 91.4%; Score 53; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 9  
 |||||

Db 1 CKKLFCKTK 9

## RESULT 5

ID R39291 standard; peptide; 9 AA.  
 AC R39291;  
 DT 22-DEC-1993 (first entry)

DE Endotoxin lipid A neutralising peptide.

KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 OS Synthetic.

FH Key Location/Qualifiers

FT disulfide bond 1. .7  
 PN W09314115-A.

PD 22-JUL-1993.

PF 14-MAY-1992; E01060.

PR 16-JAN-1992; US-819893.

PA (PORR/) PORRO M.

PI Porro M;

DR WPI; 93-243143/30.  
PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
PT forming non-toxic, antigenic complex, used to treat or prevent  
PT septic shock, in vaccines; alpha fetoprotein; homeodomain; zinc-finger.  
PS Claim 7: Page 32; 45pp; English.  
CC The sequence is that of a peptide which binds to the lipid A  
CC component of bacterial endotoxin at the same site as polymyxin B  
CC (PMB) and with about the same affinity to produce a non-toxic  
CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
CC proteolytic degradation in serum, has no antibiotic activity and no  
CC haemolytic action. It is especially used to treat or prevent septic  
CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
CC sera, vaccines, drug solns. etc.; to control release of cytokines  
CC induced by endotoxins; for in vivo or in vitro detoxification of  
CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
CC products.  
SQ Sequence 9 AA;  
  
Query Match 91.4%; Score 53; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CKKLFCKCT 9  
Db 1 CKKLFCKCT 9  
  
RESULT 6  
W21622 standard; peptide; 9 AA.  
AC W21622;  
DE 26-AUG-1997 (first entry)  
KW Antibiotic potentiating peptide #34.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..7  
PN W09638163-AL.  
PD 05-DEC-1996.  
PF 23-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence reduces dose of antibiotic required  
PS Claim 38; Page 28; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 9 AA;  
  
Query Match 91.4%; Score 53; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CKKLFCKCT 9  
Db 1 CKKLFCKCT 9  
  
RESULT 7  
R23962 standard; Protein; 2783 AA.  
ID R23962

AC R23962;  
DE 06-JAN-1993 (first entry)  
DE APP-1.  
OS Enhancer; alpha fetoprotein; homeodomain; zinc-finger.  
PH Homo sapiens.  
FH Key Location/Qualifiers  
FT domain 1231..1291  
FT /note= "homeodomain"  
FT domain 1328..1388  
FT /label= homeodomain  
FT domain 1727..1786  
FT /label= homeodomain  
FT domain 2032..2091  
FT /label= homeodomain  
FT domain 67..99  
FT /label= zinc-finger  
FT domain 123..155  
FT /label= zinc-finger  
FT domain 171..203  
FT /label= zinc-finger  
FT domain 306..366  
FT /label= zinc-finger  
FT domain 443..476  
FT /label= zinc-finger  
FT domain 484..543  
FT /label= zinc-finger  
FT domain 628..660  
FT /label= zinc-finger  
FT domain 679..711  
FT /label= zinc-finger  
FT domain 1066..1097  
FT /label= zinc-finger  
FT domain 1411..1442  
FT /label= zinc-finger  
FT domain 1613..1643  
FT /label= zinc-finger  
FT domain 1794..1825  
FT /label= zinc-finger  
FT domain 2107..2139  
FT /label= zinc-finger  
FT domain 2540..2571  
FT /label= zinc-finger  
FT domain 2606..2638  
FT /label= zinc-finger  
FT EP-487229-A.  
PN 27-MAY-1992.  
PF 07-NOV-1991; 310334.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Higashio K, Morinaga T, Tamaoki T, Yasuda N;  
DR WPI; 92-176828/22.  
DR N-PSDB; Q24828.  
PT DNA encoding protein binding to alpha-fetoprotein gene enhancer -  
PT useful for prodn. of biological active protein  
PS Claim 1; Page 7; 24pp; English.  
CC This protein specifically binds to the enhancer of the alpha-  
CC fetoprotein gene. The DNA encoding this protein may be useful for  
CC the prodn. of biologically active proteins by inserting the DNA into  
CC an expression vector and co-transfecting animal cells with another  
CC expression vector in which a gene for the biologically active protein  
CC is placed under the control of alpha-fetoprotein gene enhancer and  
CC promoter.  
SQ Sequence 2783 AA;  
  
Query Match 60.3%; Score 35; DB 1; Length 2783;  
Best Local Similarity 66.7%; Pred. No. 5.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 CKKLFCKCT 9  
Db 1802 CRALFKAKT 1810

```

RESULT 8
R23963
ID R23963 standard; Protein; 2783 AA.
AC R23963;
DT 06-JAN-1993 (first entry)
DE AFP-1 (Ala 2460 Val).
KW Enhancer; alpha fetoprotein; homeodomain; zinc-finger.
OS Homo sapiens.
FH Key
FT domain Location/Qualifiers
FT 1231..1291
FT /note= "homeodomain"
FT domain
FT 1328..1388
FT /label= homeodomain
FT domain
FT 1727..1786
FT /label= homeodomain
FT domain
FT 2032..2091
FT /label= homeodomain
FT domain
FT 67..99
FT /label= zinc-finger
FT domain
FT 123..155
FT /label= zinc-finger
FT domain
FT 171..203
FT /label= zinc-finger
FT domain
FT 306..366
FT /label= zinc-finger
FT domain
FT 443..476
FT /label= zinc-finger
FT domain
FT 484..543
FT /label= zinc-finger
FT domain
FT 628..660
FT /label= zinc-finger
FT domain
FT 679..711
FT /label= zinc-finger
FT domain
FT 1066..1097
FT /label= zinc-finger
FT domain
FT 1411..1442
FT /label= zinc-finger
FT domain
FT 1613..1643
FT /label= zinc-finger
FT domain
FT 1794..1825
FT /label= zinc-finger
FT domain
FT 2107..2139
FT /label= zinc-finger
FT domain
FT 2540..2571
FT /label= zinc-finger
FT domain
FT 2606..2638
FT /label= zinc-finger
FT domain
FT EP-487229-A.
PD 27-MAY-1992.. 310334.
PF 07-NOV-1991; JP-301412.
PR (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Higashio K, Morinaga T, Tamaoki T, Yasuda N;
PI WPI; 92-176828/22.
DR N-PSDB; 024829.
PT DNA encoding protein binding to alpha-fetoprotein gene enhancer -
PT useful for prodn. of biological active protein
PS Claim 3-4; Page 7; 24pp; English.
CC This protein sequence is as the sequence given in R23962 except
CC that Ala at position 2460 is replaced by Val.
CC This protein specifically binds to the enhancer of the alpha-
CC fetoprotein gene. The DNA encoding this protein may be useful for
CC an expression vector and co-transfecting animal cells with another
CC expression vector in which a gene for the biologically active protein
CC is placed under the control of alpha-fetoprotein gene enhancer and
CC promoter.
SQ Sequence 2783 AA;

Query Match 60.3%; Score 35; DB 1; Length 2783;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

QY 1 CKKLFCKT 9
DB 1802 CRALEFAKT 1810
I: |||||

RESULT 9
W19786
ID W19786 standard; Protein; 1088 AA.
AC W19786;
DT 08-NOV-1997 (first entry)
DE Human multiple function regulatory protein SPT5.
KW SPT5; multiple function regulatory protein; MFRP; cell cycle;
KW gastrulation; carcinogenesis; tumorigenicity; breast cancer;
KW tumour suppressor; neoplasm; carcinogenesis; organogenesis;
KW gonad development.
OS Homo sapiens.
FH Key
FT peptide Location/Qualifiers
FT 827..830
FT /note= "MAPK consensus site"
FT peptide
FT 902..905
FT /note= "MAPK consensus site"
FT peptide
FT 933..937
FT /note= "MAPK consensus site"
FT peptide
FT 944..948
FT /note= "MAPK consensus site"
FT peptide
FT 953..956
FT /note= "MAPK consensus site"
FT peptide
FT 966..969
FT /note= "MAPK consensus site"
FT peptide
FT 1033..1036
FT /note= "MAPK consensus site"
FT peptide
FT W09703087-AL.
PD 30-JAN-1997.
PF 12-JUL-1996; U11702.
PR 13-JUL-1995; US-001113.
PA (UNMI ) UNIV MICHIGAN.
PI Chiang P, Kurnit DM, Song W;
PI WPI; 97-118981/11.
DR N-PSDB; T72814.
PT New family of multiple function regulatory proteins - regulate cell
PT cycle progression, gastrulation, carcinogenesis, tumorigenicity,
PT gonad development, etc.
PS Claim 32; Fig 14b; 98pp; English.
CC This polypeptide is a novel multiple function regulatory protein
CC (MFRP) that shows homology to yeast SPT5. It is encoded by a cDNA
CC clone (T72814) obt'd. from a human lung cDNA library. Mammalian
CC SPT5, SPT4 and SPT6 (see also W19782-85) homologues of yeast SPT
CC transcription factors regulate cellular functions such as cell
CC cycle progression, control of gastrulation, carcinogenesis (such as
CC breast cancer), tumorigenicity, inhibition or reversal of a
CC neoplastic phenotype, progression through the cell cycle, gonad
CC development, carcinogenesis and organogenesis. They can be
CC expressed in transformed host cells and used to generate antibodies
CC used in a claimed method for detecting wild-type mammalian MFRP, or
CC expressed using gene therapy methods to restore wild-type MFRP
CC function, such as tumour suppression.
SQ Sequence 1088 AA;

Query Match 60.3%; Score 35; DB 1; Length 1088;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

QY 1 CKKL-----FKCKT 10
DB 627 CKKLVENGMFVCKTR 642
I: |||||

RESULT 10
R46299
ID R46299 standard; Protein; 333 AA.

```

AC R46299;  
 DT 05-SEP-1994 (first entry)  
 DE Homologue of flavonol synthase from Dianthus caryophyllus.  
 KW Flavonol synthase; plant; Dianthus caryophyllus; modulation;  
 KW transformation; manipulation; colour; production; synthesis;  
 KW dihydrokaempferol; DHK; dihydroquercetin; DHQ; dihydromyricetin;  
 KW DHM.  
 OS Dianthus caryophyllus (Clone PCGP777).  
 PN W09403606-A.  
 PD 17-FEB-1994.  
 PF 05-AUG-1993; AU0400.  
 PR 05-AUG-1992; AU-003944.  
 PA (ITEL-) INT FLOWER DEV PTY LTD.  
 PI Holton TA, Keam LA;  
 DR WPI: 94-065695/08.  
 DR N-PSDB: Q57610.  
 PT Nucleic acid encoding a plant flavonol synthase - used for  
 PT modulating flavonol prodn. in plants, for manipulating petal  
 PT colour and male fertility.  
 PS Example 11; Page 45-47; 88pp; English.  
 CC The cDNA encoding the flavonol synthase homologue can be used to  
 CC transform plants. The nucleic acid permits the modulation of levels  
 CC of flavonol production in plants, allowing the manipulation of petal  
 CC colour and male fertility. It also permits the modulation of  
 CC dihydrokaempferol (DHK) metabolism as well as the metabolism of other  
 CC substrates such as dihydroquercetin (DHQ) and dihydromyricetin (DHM),  
 CC which allows such manipulation of flower colour.  
 SQ Sequence 333 AA;

Query Match 58.6%; Score 34; DB 1; Length 333;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KLFCKTK 10  
 ||||| :|  
 Db 170 KLFCKSK 177

RESULT 11  
 R06646  
 ID R06646 standard; protein; 551 AA.  
 AC R06646;  
 DT 07-JAN-1991 (first entry)  
 DE Human interleukin-2 receptor beta-chain.  
 KW IL-2R beta; pIL-2R beta 9.  
 OS Homo sapiens.  
 PN EP-386304-A.  
 PD 12-SEP-1990.  
 PF 29-MAY-1989; 109656.  
 PR 07-MAR-1989; EP-104023.  
 PR 29-MAY-1989; EP-109656.  
 PA (OSAU ) OSAKA UNIVERSITY.  
 PI Taniguchi T;  
 DR WPI: 90-276456/37.  
 DR N-PSDB: Q05872.  
 PT Recombinant interleukin-2 receptor beta chain - used for studying  
 PT IL-2 system and producing antibodies for diagnosis and therapy.  
 PS Claim 4; Fig 1B; 37pp; English.  
 CC IL-2 receptor beta chains are useful in studying the biochemistry of  
 CC interleukin, and in diagnosis and therapy by immune suppression and  
 CC activation.  
 SQ Sequence 551 AA;

Query Match 56.9%; Score 33; DB 1; Length 551;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KLFCKCKT 9  
 ||||| :|  
 Db 275 KLLKCKNT 282

RESULT 12  
 R07506  
 ID R07506 standard; protein; 551 AA.  
 AC R07506;  
 DT 07-FEB-1991 (first entry)  
 DE IL-2R beta chain.  
 KW Interleukin; receptor.  
 OS Homo sapiens.  
 FH Key  
 FT peptide  
 FT 1..26  
 FT /label=signal peptide  
 FT 27..551  
 FT /label=sol.IL-2R beta  
 FT 241..265  
 FT domain  
 FT /label=transmembrane region  
 FT 29..31  
 FT /label=N-glycos\_site  
 FT 43..45  
 FT modified\_site  
 FT /label=N-glycos\_site  
 FT 71..73  
 FT modified\_site  
 FT /label=N-glycos\_site  
 FT 149..151  
 FT modified\_site  
 FT /label=N-glycos\_site  
 FT 352..354  
 FT modified\_site  
 FT /label=N-glycos\_site  
 PN AU050726-A.  
 PD 13-SEP-1990.  
 PF 06-MAR-1990; 050726.  
 PR 07-MAR-1989; EP-104023.  
 PR 29-MAY-1989; EP-109656.  
 PR 20-JUL-1989; EP-113310.  
 PA (BOEH ) BOEHRINGER INGELHEI.  
 PI Taniguchi T, Hatakeyama M, Minamoto S, Kono T, Doi T;  
 PI Miyazaka M, Tsudo, Karasuyama H;  
 DR WPI: 90-327673/44.  
 DR N-PSDB: Q06331.  
 PT Recombinant interleukin-2 receptor beta chain - useful for  
 PT diagnosis and therapy by immune suppression or activation.  
 PS Claim 3; Fig 1B; 65pp; English.  
 CC The sequence was deduced from cDNA obtd. from clones isolated from  
 CC a cDNA library prepd. from RNA extracted from the NK-like human  
 CC lymphoid cell line, Yt. The clones were selected using a cocktail  
 CC of anti-IL-2R beta MAbs, Mik-beta1 and Mik-beta2. The sequence  
 CC encodes a soluble portion of the IL-2R beta chain.  
 CC See also R07507.  
 SQ Sequence 551 AA;

Query Match 56.9%; Score 33; DB 1; Length 551;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KLFCKCKT 9  
 ||||| :|  
 Db 275 KLLKCKNT 282

RESULT 13  
 W88480  
 ID W88480 standard; protein; 971 AA.  
 AC W88480;  
 DT 10-MAY-1999 (first entry)  
 DE Candida albicans histidine kinase 1 CAHk1 partial sequence.  
 KW CAHk1 gene; histidine kinase-1; two-component gene;  
 KW signal transduction; virulence; candidosis; diagnosis; therapy;  
 KW vaccine; infection.  
 OS Candida albicans.  
 FH Key  
 FT Location/Qualifiers  
 FT 482..721  
 FT /note= sensor domain, this polypeptide is  
 FT specifically claimed in Claim 13(g)  
 FT Domain 834..971

FT FT /note= "response regulator domain, this polypeptide  
 FT FT is specifically claimed in Claim 13(h)"  
 FT FT 2. .971  
 FT FT /note= "this polypeptide region is specifically  
 FT FT claimed in Claim 13(c)"  
 FT FT 482. .971  
 FT FT /note= "this polypeptide region is specifically  
 FT FT claimed in Claim 13(i)"  
 PN WO9902700-A1.  
 PD 21-JAN-1999.  
 PF 09-JUL-1998; U14254.  
 PR 11-FEB-1998; US-074308.  
 PR 10-JUL-1997; US-052273.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (GEOU) UNIV GEORGETOWN.  
 PI Calderone RA, Calera JA, Choi GH;  
 DR WPI; 99-120901/10.  
 DR N-PSDB; X06981.  
 PT New isolated Candida albicans Histidine Kinase-1 gene - used to  
 PT develop products for the diagnosis, prevention and treatment of C.  
 PT albicans infections, particularly oral candidosis  
 PS Claim 13(a): Page 86-89; 106pp; English.  
 CC This polypeptide comprises a portion of Candida albicans  
 CC histidine kinase-1 (CaHk1), and is encoded by a partial length  
 CC DNA clone (see X06981) deposited as ATCC 209504. The full-length  
 CC 2471-amino acid CaHk-1 protein (282 kDa) is provided in W88481.  
 CC CaHk1 is active in phosphorylating host cell proteins to render  
 CC the host cell susceptible to invasion by C. albicans, i.e. is  
 CC involved in virulence. It exhibits homology with the sensor and  
 CC regulator components of prokaryotic and eukaryotic two-component  
 CC histidine kinases. Also provided are vectors, host cells,  
 CC antibodies and recombinant methods for producing them. The  
 CC invention also provides agonists and antagonists of CaHk1 activity,  
 CC as well as diagnostic methods for detecting CaHk1 nucleic acids,  
 CC polypeptides and antibodies in a biological sample. CaHk1  
 CC antagonists can be used to treat C. albicans infections, e.g. oral  
 CC candidosis, including oral candidosis associated with HIV  
 CC infection. The products can also be used in e.g. immunoassays, in  
 CC epitope mapping, epitope tagging, to generate antibodies to a  
 CC particular portion of the protein, and as vaccines.  
 SQ Sequence 971 AA;

Query Match 56.9%; Score 33; DB 1; Length 971;  
 Best Local Similarity 55.6%; Pred. NO. 4.5e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKKT 9  
 DB 7 CSQHFCKKS 15

RESULT 14

ID R92103 standard; Protein; 1706 AA.  
 AC R92103;  
 DT 19-MAY-1996 (first entry)  
 DE Rat RIZ.  
 KW Retinoblastoma protein-interacting zinc finger; RIZ;  
 KW cell proliferation; tumour; cancer; neuroblastoma; melanoma;  
 KW diagnosis; therapy.  
 OS Rattus sp.  
 FH Key Location/Qualifiers  
 FT domain 36. .151  
 FT domain /label= PR\_domain  
 FT domain 304. .309  
 FT region /label= cr2\_core\_motif  
 FT region 357. .377  
 FT region /label= zinc\_finger(1)  
 FT region 387. .407  
 FT region /label= zinc\_finger(2)  
 FT region 478. .499  
 FT region /label= zinc\_finger(3)

FT region 667. .695  
 FT domain /label= Leucine-zipper  
 FT domain 867. .874  
 FT region /label= Nuclear\_localisation\_signal  
 FT region 1125. .1203  
 FT region /label= Zinc\_finger(4-6)  
 FT region 1323. .1343  
 FT region /label= Zinc\_finger(7)  
 FT region 1445. .1466  
 FT region /label= Zinc\_finger(8)  
 PN WO9606168-A2.  
 PD 29-FEB-1996.  
 PE 18-AUG-1995; U10574.  
 PR 18-AUG-1994; US-292683.  
 PR 06-MAR-1995; US-399411.  
 PA (LJOL-) LA JOLLA CANCER RES FOUND.  
 PI Huang S;  
 DR WPI; 96-151371/15.  
 DR N-PSDB; T18023.  
 PT Nucleic acid encoding mammalian retinoblastoma protein-interacting  
 PT zinc finger - used to identify cell proliferation modulating agents  
 PT for treatment of tumours, esp. neuroblastoma and melanoma, also used  
 PT for cancer diagnosis.  
 PS Claim 11: Fig 1A; 142pp; English.  
 CC The sequence (R92103) of full-length rat retinoblastoma (Rb)  
 CC protein-interacting zinc finger (RIZ) was deduced from a cDNA  
 CC sequence (T18023) cloned from a rat neonatal cardiac myocyte  
 CC library. RIZ is expressed in rat cells as a 250 kDa  
 CC phosphoprotein. It acts as a cell differentiation factor, and  
 CC can modulate cell function by binding to Rb, which is involved in  
 CC regulating cell proliferation. RIZ can also act to regulate  
 CC transcription. RIZ or its active fragments (see R92105) can be  
 CC obt'd. by recombinant DNA methods, and used to identify agents that  
 CC modulate cell proliferation and thereby treat tumour growth or  
 CC neurodegenerative disorders.  
 SQ Sequence 1706 AA;

Query Match 55.2%; Score 32; DB 1; Length 1706;  
 Best Local Similarity 66.7%; Pred. NO. 1.1e+03;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKKT 9  
 DB 1156 CVOLEFKVKT 1164

RESULT 15

W4777  
 ID W4777 standard; Protein; 1141 AA.  
 AC W4777;  
 DT 01-JUN-1998 (first entry)  
 DE Human Tbc-1 protein.  
 KW tre-2; BUB2; cdc16; mast cell; probe; hybridisation; subtraction method;  
 KW mouse; transcription factor; differentiation; proliferation; human;  
 KW acute myelogenous leukaemia.  
 OS Homo sapiens.  
 PN US5700927-A.  
 PD 23-DEC-1997.  
 PR 23-DEC-1994; 363300.  
 PR 23-DEC-1994; US-363300.  
 PA (CHIL-) CHILDRENS MEDICAL CENT.  
 PI Richardson P, Zon L;  
 DR WPI; 98-062437/06.  
 DR N-PSDB; V05886.  
 PT DNA encoding Tbc1 polypeptide - useful for treating leukaemia  
 PS Claim 1; Fig 1A-B; 22pp; English.  
 CC This is the amino acid sequence of a Tbc1 (tre-2, BUB2 and cdc16)  
 CC polypeptide, isolated from a bone marrow-derived mast cell DNA library.  
 CC The screen was carried out using a probe generated by a subtraction  
 CC method which compared mRNA expression in an undifferentiated mast cell  
 CC line P815 and cell line PGT6 (P815 cells transfected to express the  
 CC murine GATA-1 transcription factor - a factor which controls the

CC expression of genes involved in mast cell differentiation). Tbc1 encodes  
CC a protein involved in the coupling of cell proliferation to cell  
CC differentiation, which can be used to treat leukaemia (especially acute  
CC myelogenous leukaemia) by causing leukaemic cells to differentiate.  
SQ Sequence 1141 AA;

Query Match 55.2%; Score 32; DB 1; Length 1141;  
Best Local Similarity 50.0%; Pred. No. 7.5e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCK 8  
| :|||  
Db 54 CSSIFECK 61

Search completed: September 7, 1999, 22:50:05  
Job time: 7886 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:49 ; Search time 82.37 Seconds  
(without alignments)  
1.198 Million cell updates/sec

Title: US-09-124-280A-33  
Perfect score: 58  
Sequence: 1 CKKLFCKTK 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	10	1	US-08-049-871-4
2	58	100.0	10	1	US-07-819-893-4
3	58	100.0	10	1	US-08-280-397-4
4	58	100.0	10	2	US-08-456-112B-33
5	53	91.4	9	1	US-08-049-871-5
6	53	91.4	9	1	US-07-819-893-5
7	53	91.4	9	1	US-08-280-397-5
8	53	91.4	9	2	US-08-456-112B-34
9	34	58.6	333	1	US-08-379-556A-8
10	32	55.2	1141	1	US-08-363-300-2
11	32	55.2	1706	2	US-08-459-568-2
12	32	55.2	86	2	US-08-459-568-82
13	32	55.2	1706	2	US-08-399-411-2
14	32	55.2	86	2	US-08-399-411-82
15	31.5	54.3	10	1	US-08-097-830E-18
16	31.5	54.3	10	2	US-08-456-112B-18
17	31	53.4	122	1	US-07-734-534A-1
18	31	53.4	200	4	518960-8
19	30	51.7	32	1	US-08-361-920-10
20	30	51.7	514	1	US-08-361-920-21
21	30	51.7	740	1	US-08-276-099A-12
22	30	51.7	32	1	US-08-479-939-10
23	30	51.7	514	1	US-08-479-939-21
24	30	51.7	740	1	US-08-781-890-12
25	30	51.7	750	1	US-08-369-796-4
26	30	51.7	712	1	US-08-369-796-6
27	30	51.7	749	1	US-08-369-796-8
28	30	51.7	32	2	US-08-483-432-10
29	30	51.7	514	2	US-08-483-432-21
30	30	51.7	1719	2	US-08-459-568-4
31	30	51.7	1719	2	US-08-399-411-4
32	30	51.7	11	2	US-08-456-112B-13
33	30	51.7	750	2	US-08-852-091-4
34	30	51.7	712	2	US-08-852-091-6
35	30	51.7	749	2	US-08-852-091-8
36	30	51.7	750	3	PCT-US95-17025-4
37	30	51.7	712	3	PCT-US95-17025-6
38	30	51.7	749	3	PCT-US95-17025-8
39	29.5	50.9	374	1	US-08-450-393A-2

Sequence 4, Appl1  
Sequence 3, Appl1  
Sequence 2, Appl1  
Sequence 4, Appl1  
Sequence 9, Appl1  
Sequence 54, Appl1

40 29.5 50.9 360 1 US-08-450-393A-4  
41 29.5 50.9 347 2 US-08-461-244-3  
42 29.5 50.9 374 3 PCT-US95-00476-2  
43 29.5 50.9 360 3 PCT-US95-00476-4  
44 29 50.0 133 1 US-07-893-929A-9  
45 29 50.0 1141 1 US-08-131-365B-54

ALIGNMENTS

RESULT 1  
US-08-049-871-4  
; Sequence 4, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-049-871-4

Query Match 100.0%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00035; Indels 0; Caps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 1 CKKLFCKTK 10  
| | | | | | | | | |  
Db 1 CKKLFCKTK 10

RESULT 2  
US-07-819-893-4  
; Sequence 4, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the

;  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
; US-07-819-893-4

Query Match 100.0%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00035;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFKCKTK 10  
Db 1 CKKLFKCKTK 10

RESULT 3  
US-08-280-397-4  
; Sequence 4, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porto, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514

;  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-280-397-4

Query Match 100.0%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00035;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFKCKTK 10  
Db 1 CKKLFKCKTK 10

RESULT 4  
US-08-456-112B-33  
; Sequence 33, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porto, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-33

Query Match 100.0%; Score 58; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00035;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10  
| | | | |  
Db 1 CKKLFCKTK 10

RESULT 5  
US-08-049-871-5  
; Sequence 5, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-5

Query Match 91.4%; Score 53; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 9  
| | | | |  
Db 1 CKKLFCKTK 9

RESULT 6  
US-07-819-893-5  
; Sequence 5, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas

; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
US-07-819-893-5

Query Match 91.4%; Score 53; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 9  
| | | | |  
Db 1 CKKLFCKTK 9

RESULT 7  
US-08-280-397-5  
; Sequence 5, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-5

Query Match 91.4%; Score 53; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKT 9  
Db 1 CKKLFCKT 9

RESULT 8  
US-08-456-112B-34  
Sequence 34, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: PORTO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-34

Query Match 91.4%; Score 53; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKT 9  
Db 1 CKKLFCKT 9

RESULT 9

US-08-379-556A-8  
Sequence 8, Application US/08379556A  
Patent No. 5859329  
GENERAL INFORMATION:  
APPLICANT: HOLTON, TIMOTHY A.  
APPLICANT: KEAM, LISA A.  
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL  
TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER  
STREET: 400 GARDEN CITY PLAZA  
CITY: GARDEN CITY  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 11530

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/379,556A  
FILING DATE: 22-MAR-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: DIGIGLIO, FRANK S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9592  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516)742-4343  
TELEFAX: (516)742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 333 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-379-556A-8

Query Match 58.6%; Score 34; DB 2; Length 333;  
Best Local Similarity 75.0%; Pred. No. 66;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KLFCKTK 10  
Db 170 KLFCKMSK 177

RESULT 10  
US-08-363-300-2  
Sequence 2, Application US/08363300  
Patent No. 5700927  
GENERAL INFORMATION:  
APPLICANT: ZOP, Leonard and Richardson, Paul  
TITLE OF INVENTION: Tdcl Gene and Uses Thereof  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/363,300

;; FILING DATE:  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Clark, Paul T.  
;; REGISTRATION NUMBER: 30,162  
;; REFERENCE/DOCKET NUMBER: 04590/009001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 542-5070  
;; TELEFAX: (617) 542-8906  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1141 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-363-300-2

Query Match 55.2%; Score 32; DB 1; Length 1141;  
Best Local Similarity 50.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKKLFCKK 8  
| :||| |  
Db 54 CSSIFECK 61

RESULT 11  
US-08-459-568-2  
; Sequence 2, Application US/08459568  
; Patent No. 5811304  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; TITLE OF INVENTION: Zinc Finger Proteins  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,568  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/399,411  
; FILING DATE: 06-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1264  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9901  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1706 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-459-568-2

Query Match 55.2%; Score 32; DB 2; Length 1706;  
Best Local Similarity 66.7%; Pred. No. 6.2e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CKKLFCKKT 9  
| :||| |  
Db 1156 CVQLFKVKT 1164

RESULT 12  
US-08-459-568-82  
; Sequence 82, Application US/08459568  
; Patent No. 5811304  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; TITLE OF INVENTION: Zinc Finger Proteins  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,568  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/399,411  
; FILING DATE: 06-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1264  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9901  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 82:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-459-568-82

Query Match 55.2%; Score 32; DB 2; Length 86;  
Best Local Similarity 66.7%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKKLFCKKT 9  
| :||| |  
Db 32 CVQLFKVKT 40

RESULT 13  
US-08-399-411-2  
; Sequence 2, Application US/08399411  
; Patent No. 5831008  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; TITLE OF INVENTION: Zinc Finger Proteins  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California

COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/399,411  
FILING DATE: 06-MAR-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1706 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-399-411-2

Query Match 55.2%; Score 32; DB 2; Length 1706;  
Best Local Similarity 66.7%; Pred. No. 6.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKKT 9  
| :||| ||  
DB 1156 CVQLFKVKKT 1164

RESULT 14  
US-08-399-411-82  
; Sequence 82, Application US/08399411  
; Patent No. 5831008  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; TITLE OF INVENTION: Zinc Finger Proteins  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/399,411  
; FILING DATE: 06-MAR-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1264  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 82:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

US-08-399-411-82

Query Match 55.2%; Score 32; DB 2; Length 86;  
Best Local Similarity 66.7%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKKT 9  
| :||| ||  
DB 32 CVQLFKVKKT 40

RESULT 15  
US-08-097-830E-18  
; Sequence 18, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-18

Query Match 54.3%; Score 31.5; DB 1; Length 10;  
Best Local Similarity 77.8%; Pred. No. 6.7;  
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CK-KLFCKK 8  
| :||| ||  
DB 2 CKFKFKCK 10

Search completed: September 7, 1999, 23:07:49  
Job time: 1747 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.  
OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:47 ; Search time 105.14 Seconds  
(without alignments)  
3.811 Million cell updates/sec

Title: US-09-124-280A-33  
Perfect score: 58  
Sequence: 1 CKKLFCKYK 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	74.1	414	2 J01016	glycine hydroxymet
2	39	67.2	32	2 S29217	neurotoxin Tx2 - s
3	39	67.2	154	2 D39384	finger protein HTF
4	39	67.2	803	2 S26823	finger protein ZNF
5	38	65.5	241	2 E70470	conserved hypotet
6	37	63.8	527	2 S53835	NADH dehydrogenase
7	37	63.8	402	2 A69002	glutamyl-tRNA redu
8	37	63.8	1051	2 A35763	collagen alpha 2 c
9	37	63.8	3198	2 A43426	collagen alpha 2 f
10	37	63.8	328	2 G65110	hypothetical 35.2
11	37	63.8	318	2 S51261	probable beta-glyc
12	36	62.1	704	2 F64373	hypothetical prote
13	35	60.3	2783	1 A41948	alpha-fetoprotein
14	35	60.3	668	2 C24785	hypothetical prote
15	35	60.3	204	2 S07741	hypothetical prote
16	35	60.3	2918	2 A54105	fibrillin-2 precur
17	35	60.3	2907	2 A57278	fibrillin-2 precur
18	35	60.3	576	2 A48157	renal transcriptio
19	34	58.6	420	2 C71634	serine hydroxymet
20	34	58.6	118	2 C64441	conserved hypotet
21	34	58.6	424	2 S11676	spore coat protein
22	34	58.6	116	2 S50449	hypothetical prote
23	34	58.6	770	2 S69075	probable membrane
24	34	58.6	86	2 I61885	zinc finger protei
25	34	58.6	189	2 F39384	finger protein HTF
26	34	58.6	163	2 G39384	finger protein HTF
27	34	58.6	1191	2 S35305	zinc finger protei
28	33	56.9	1130	1 TVHUA	protein-tyrosine k
29	33	56.9	522	1 KJHUG6	N-acetylglactosam
30	33	56.9	568	1 RRN239	polymerase-associa
31	33	56.9	568	1 RRN235	polymerase-associa
32	33	56.9	568	1 RRN273	polymerase-associa
33	33	56.9	568	1 RRN283	polymerase-associa
34	33	56.9	1123	2 A39662	kinase-related tra
35	33	56.9	397	2 S42549	polysaccharuronase
36	33	56.9	412	2 B71718	aspartate aminotra
37	33	56.9	100	2 B40651	hypothetical prote
38	33	56.9	399	2 T00631	hypothetical prote
39	33	56.9	3002	2 A47221	fibrillin 1 precur

fibrillin I - bovi  
fibrillin-1 precur  
membrane protein 4  
Na+/sulfate cotran  
NADH dehydrogenase  
hypothetical aden1

ALIGNMENTS

RESULT 1  
JQ1016  
glycine hydroxymethyltransferase (EC 2.1.2.1) - Campylobacter jejuni  
N;Alternate names: serine aldolase; serine hydroxymethylase; serine methylase; threon  
C;Species: Campylobacter jejuni  
C;Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 08-Sep-1997  
C;Accession: JQ1016; S16468  
R;Chan, V.L.; Bingham, H.L.  
Gene 101, 51-58, 1991  
A;Title: Complete sequence of the Campylobacter jejuni glyA gene encoding serine hydr  
A;Reference number: JQ1016; MUID:91285434  
A;Accession: JQ1016  
A;Molecule type: DNA  
A;Residues: 1-414 <CHAL>  
A;Cross-references: EMBL:X53816; NID:g40533; PID:g40534  
R;Chan, V.L.; Bingham, H.L.  
Gene 101, 51-58, 1990  
A;Title: Complete sequence of the Campylobacter jejuni glyA gene encoding serine hydr  
A;Reference number: S16468  
A;Accession: S16468  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-414 <CHAZ>  
A;Cross-references: EMBL:X53816; NID:g40533; PID:g40534  
C;Comments: In the presence of tetrahydrofolate, this enzyme catalyzes the reversible  
C;Gene: glyA  
C;Superfamily: glycine hydroxymethyltransferase  
C;Keywords: transferase

Query Match 74.1%; Score 43; DB 2; Length 414;  
Best Local Similarity 87.5%; Pred. No. 2.3;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKKLFCKK 8  
Db 77 CKKLFCKK 84

RESULT 2  
S29217  
neurotoxin Tx2 - spider (Phoneutria nigriventer)  
C;Species: Phoneutria nigriventer  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Jul-1997  
C;Accession: S29217  
R;do Nascimento Cordeiro, M.; Ribeiro Diniz, C.; do Carmo Valentim, A.; von Eickstedt  
FEBS Lett. 310, 153-156, 1992  
A;Title: The purification and amino acid sequences of four Tx2 neurotoxins from the v  
A;Reference number: S29214  
A;Accession: S29217  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-32 <COR>

Query Match 67.2%; Score 39; DB 2; Length 32;  
Best Local Similarity 77.8%; Pred. No. 1.4;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKKLFCKCKT 9  
Db 11 CKKLFCKCKT 9

Db 15 CCKFKCKT 23

## RESULT 3

D39384

finger protein HTF6 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 10-Sep-1997

C:Accession: D39384

R:Bellefroid, E.J.; Poncellet, D.A.; Lecocq, P.J.; Revelant, O.; Martial, J.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 3608-3612, 1991

A:Title: The evolutionarily conserved Krueppel-associated box domain defines a subfamily

A:Reference number: A39384; MUID:91219421

A:Accession: D39384

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-154 <BEL>

A:Cross-references: GB:M61869; NID:g184449; PID:g184450

C:Keywords: DNA binding; zinc; zinc finger

Query Match 67.2%; Score 39; DB 2; Length 154;

Best Local Similarity 100.0%; Pred. No. 4.9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLFKCK 8

|||||

Db 129 KKLFKCK 135

## RESULT 4

S26823

finger protein ZNF43 - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Sep-1998

C:Accession: S26823

R:Lovering, R.; Trowsdale, J.

Nucleic Acids Res. 19, 2921-2928, 1991

A:Title: A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell li

A:Reference number: S26823; MUID:91279444

A:Accession: S26823

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-803 <LOW>

A:Cross-references: EMBL:X59244; NID:g38031; PID:g38032

C:Genetics:

A:Gene: GDB:2NF43; HTF6

A:Cross-references: GDB:128653

A:Map position: 19p13.1-19p12

C:Keywords: DNA binding

Query Match 67.2%; Score 39; DB 2; Length 803;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLFKCK 8

|||||

Db 165 KKLFKCK 171

## RESULT 5

E70470

conserved hypothetical protein aq\_1986 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 08-May-1998

C:Accession: E70470

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

V. Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666

A:Accession: E70470

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-241 <AQE>

A:Cross-references: GB:AE000766; NID:g2984216; PID:g2984233; GB:AE000657

A:Experimental source: strain VF5

C:Genetics:

A:Gene: aq\_1986

Query Match 65.5%; Score 38; DB 2; Length 241;

Best Local Similarity 75.0%; Pred. No. 10;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCKLFCKCK 8

|||||

Db 162 CEKLLKCK 169

## RESULT 6

S53835

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - Acanthamoeba castellanii mitoc

C:Species: mitochondrion Acanthamoeba castellanii

C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Feb-1999

C:Accession: S53835

R:Burger, G.; Plante, I.; Lonergan, K.M.; Gray, M.W.

J. Mol. Biol. 245, 522-537, 1995

A:Title: The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii: c

A:Reference number: S53825; MUID:95147275

A:Accession: S53835

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-527 <BUR>

A:Cross-references: GB:U12386; NID:g562028; PID:g562039

A:Experimental source: strain Neff; ATCC 30010

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC6

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 2

C:Keywords: mitochondrion; NAD; oxidoreductase

Query Match 63.8%; Score 37; DB 2; Length 527;

Best Local Similarity 87.5%; Pred. No. 29;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKLFKCKT 9

|||||

Db 13 KKLFKCKT 20

## RESULT 7

A69002

glutamyl-tRNA reductase - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998

C:Accession: A69002

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivani,

Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A:Reference number: A69000; MUID:98037514

A:Accession: A69002

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-402 <MTH>

A:Cross-references: GB:AE000874; GB:AE000666; NID:g2622110; PID:g2622114

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1012

A:Start codon: GTG

C:Superfamily: glutamyl-tRNA reductase



C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 29-Jan-1999

C;Accession: G65110  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of *Escherichia coli* K-12.

A;Reference number: A64720; MUID:97426617

A;Accession: G65110

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-328 <BLAT>

A;Cross-references: GB:AE000399; GB:U00096; NID:92367201; PID:gl789588; UWGP:b3197

A;Experimental source: strain K-12, substrain MGL655

C;Genetics:

A;Gene: yrbH

C;Superfamily: probable ATP-binding protein gutQ; CBS homology

Query Match 63.8%; Score 37; DB 2; Length 328;

Best Local Similarity 60.0%; Pred. No. 20;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10

||:|:|:|

DB 42 CERMFCKGK 51

RESULT 11

S51261

probable beta-glycosyltransferase trsB - *Yersinia enterocolitica*

C;Species: *Yersinia enterocolitica*

C;Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999

C;Accession: S51261; S70737

R;Skurnik, M.; Venho, R.; Toivanen, P.; Al-Hendy, A.

submitted to the EMBL Data Library, January 1995.

A;Description: A novel locus of *Yersinia enterocolitica* serotype O:3 involved in lipopolys

A;Reference number: S51258

A;Accession: S51261

A;Molecule type: DNA

A;Residues: 1-318 <SKU>

A;Cross-references: EMBL:247767; NID:9633689; PID:9633693

A;Experimental source: strain 6471/76 serotype O:3

R;Skurnik, M.; Venho, R.; Toivanen, P.; Al-Hendy, A.

Mol. Microbiol. 17, 575-594, 1995

A;Title: A novel locus of *Yersinia enterocolitica* serotype O:3 involved in lipopolys

A;Reference number: S70734; MUID:96100456

A;Accession: S70737

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-110 <SKW>

A;Cross-references: EMBL:247767

A;Experimental source: strain 6471/76 serotype O:3

C;Genetics:

A;Gene: trsB

C;Keywords: lipopolysaccharide core biosynthesis

Query Match 63.8%; Score 37; DB 2; Length 318;

Best Local Similarity 62.5%; Pred. No. 19;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCKK 8

||:|:|:|

DB 193 CKSFECK 200

RESULT 12

F64373

hypothetical protein homolog MJ0590 - *Methanococcus jannaschii*

C;Species: *Methanococcus jannaschii*

C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1998

C;Accession: F64373

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

Query Match 63.8%; Score 37; DB 2; Length 402;

Best Local Similarity 75.0%; Pred. No. 23;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKK 8

|||||

DB 392 CAKLFCK 399

RESULT 8

A35763

collagen alpha 2 chain - sea urchin (*Paracentrotus lividus*) (fragment)

C;Species: *Paracentrotus lividus* (common urchin)

C;Date: 12-Oct-1990 #sequence\_revision 12-Oct-1990 #text\_change 08-Sep-1997

C;Accession: A35763

R;D'Alessio, M.; Ramirez, F.; Suzuki, H.R.; Solursh, M.; Gambino, R.

J. Biol. Chem. 265, 7050-7054, 1990

A;Title: Cloning of a fibrillar collagen gene expressed in the mesenchymal cells of the

A;Reference number: A35763; MUID:90216744

A;Accession: A35763

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1051 <DAA>

A;Cross-references: GB:J05422; NID:9159961; PID:gl59962

C;Superfamily: unassigned collagens; fibrillar collagen carboxyl-terminal homology

F:830-1051/Domain: fibrillar collagen carboxyl-terminal homology <RCC>

Query Match 63.8%; Score 37; DB 2; Length 1051;

Best Local Similarity 85.7%; Pred. No. 50;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCK 7

|||||

DB 854 CKDLFKC 860

RESULT 9

A43426

collagen alpha 2 fibrillar chain precursor - sea urchin (*Strongylocentrotus purpuratus*)

C;Species: *Strongylocentrotus purpuratus* (purple urchin)

C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998

C;Accession: A43426

R;Exposito, J.Y.; D'Alessio, M.; Ramirez, F.

J. Biol. Chem. 267, 17404-17408, 1992

A;Title: Novel amino-terminal propeptide configuration in a fibrillar procollagen underg

A;Reference number: A43426; MUID:92381062

A;Accession: A43426

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-3198 <EXP>

A;Cross-references: GB:M92041; NID:9161448; PID:9161449

A;Note: sequence extracted from NCBI backbone (NCBIP:111965)

C;Superfamily: unassigned collagens; fibrillar collagen carboxyl-terminal homology; von

F:48-106/Domain: von Willebrand factor type C repeat homology <VWC>

F:2978-3198/Domain: fibrillar collagen carboxyl-terminal homology <RCC>

Query Match 63.8%; Score 37; DB 2; Length 3198;

Best Local Similarity 85.7%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCK 7

|||||

DB 3002 CKDLFKC 3008

RESULT 10

G65110

hypothetical 35.2 kD protein in murZ-rpoN intergenic region - *Escherichia coli* (strain K

C;Species: *Escherichia coli*

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A:Reference number: A64300; MUID:96337999  
 A:Accession: F64373  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-704 <BUL>  
 A:Cross-references: GB:U67507; GB:L77117; NID:g1591288; PID:g1591298; TIGR:MJ0590; PID:g  
 C:Genetics:  
 A:Map position: FOR522030-524144  
 C:Superfamily: hypothetical protein MJ0590

Query Match 62.1%; Score 35; DB 2; Length 704;  
 Best Local Similarity 70.0%; Pred. No. 54;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10  
 |||||  
 DB 521 CKKLGKCVNK 530

## RESULT 13

A41948  
 alpha-fetoprotein enhancer-binding protein - human  
 N:Alternate names: ATF1 protein  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 14-Nov-1997  
 C:Accession: A41948  
 R:Morinaga, T.; Yasuda, H.; Hashimoto, T.; Higashio, K.; Tamaoki, T.  
 Mol. Cell. Biol. 11, 6041-6049, 1991  
 A:Title: A human alpha-fetoprotein enhancer-binding protein, ATF1, contains four homeod  
 A:Reference number: A41948; MUID:92049333  
 A:Accession: A41948  
 A:Molecule type: mRNA  
 A:Residues: 1-2783 <MOR>  
 A:Cross-references: GB:D10250; GB:D90395; NID:g219429; PID:g219430  
 A:Note: sequence extracted from NCBI backbone (NCBI:66271, NCBI:66276)  
 C:Genetics:  
 A:Gene: GDB:ATBF1  
 A:Cross-references: GDB:392090; OMIM:104155  
 A:Map position: 16q22.3-16q23.1  
 C:Superfamily: alpha-fetoprotein enhancer-binding protein; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription factor; zinc finger  
 F:72-94/Region: zinc finger CCHH motif  
 F:128-150/Region: zinc finger CCHH motif  
 F:176-198/Region: zinc finger CCHH motif  
 F:311-332/Region: zinc finger CCHH motif  
 F:340-361/Region: zinc finger CCHH motif  
 F:448-471/Region: zinc finger CCHH motif  
 F:489-509/Region: zinc finger CCHH motif  
 F:517-538/Region: zinc finger CCHH motif  
 F:633-655/Region: zinc finger CCHH motif  
 F:684-706/Region: zinc finger CCHH motif  
 F:719-773/Region: zinc finger CCHH motif  
 F:809-958/Region: glutamine-rich  
 F:1071-1092/Region: zinc finger CCHH motif  
 F:1117-1211/Region: proline-rich  
 F:1332-1288/Domain: homeobox homology <HOX1>  
 F:1329-1385/Domain: homeobox homology <HOX2>  
 F:1416-1437/Region: zinc finger CCHH motif  
 F:1618-1638/Region: zinc finger CCHH motif  
 F:1728-1784/Domain: homeobox homology <HOX3>  
 F:1799-1820/Region: zinc finger CCHH motif  
 F:2033-2089/Domain: homeobox homology <HOX4>  
 F:2112-2134/Region: zinc finger CCHH motif  
 F:2545-2566/Region: zinc finger CCHH motif  
 F:2585-2607/Region: glycine-rich  
 F:2611-2633/Region: zinc finger CCHH motif  
 F:2650-2737/Region: serine/threonine-rich

Query Match 60.3%; Score 35; DB 1; Length 2783;  
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 9  
 |||||  
 DB 1802 CRALFKAKT 1810

## RESULT 14

C24785  
 hypothetical protein 668 - slime mold (Dictyostelium discoideum) transposon DIRS-1 (f  
 C:Species: Dictyostelium discoideum  
 C:Date: 23-Aug-1987 #sequence\_revision 23-Aug-1987 #text\_change 18-Jun-1993  
 C:Accession: C24785  
 R:Cappello, J.; Handelsman, K.; Lodish, H.F.  
 Cell 43, 105-115, 1985  
 A:Title: Sequence of Dictyostelium DIRS-1: an apparent retrotransposon with inverted  
 A:Reference number: A94654; MUID:86079481  
 A:Accession: C24785  
 A:Molecule type: DNA  
 A:Residues: 1-668 <CAP>

Query Match 60.3%; Score 35; DB 2; Length 668;  
 Best Local Similarity 60.0%; Pred. No. 77;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10  
 |||||  
 DB 440 CRKLNCKLK 449

## RESULT 15

S07741  
 hypothetical protein 9 - Paramecium tetraurelia mitochondrion (SGC6)  
 C:Species: mitochondrion Paramecium tetraurelia  
 C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 09-Sep-1997  
 C:Accession: S07741  
 R:Pritchard, A.E.; Seilhamer, J.J.; Mahalingam, R.; Sable, C.L.; Venuti, S.E.; Cummin  
 Nucleic Acids Res. 18, 173-180, 1990  
 A:Title: Nucleotide sequence of the mitochondrial genome of Paramecium.  
 A:Reference number: S07725; MUID:90174913  
 A:Accession: S07741  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-204 <PRI>  
 A:Cross-references: EMBL:X15917; NID:g13256; PID:g578758  
 C:Genetics:  
 A:Genome: mitochondrion  
 A:Genetic code: SGC6  
 A:Start codon: ATC  
 C:Keywords: mitochondrion

Query Match 60.3%; Score 35; DB 2; Length 204;  
 Best Local Similarity 60.0%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10  
 |||||  
 DB 114 CKKFFKLKR 123

Search completed: September 7, 1999, 23:22:49  
 Job time: 956 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:13 ; Search time 71.09 Seconds  
(without alignments)  
3.976 Million cell updates/sec

Title: US-09-124-280A-33

Perfect score: 58

Sequence: 1 CKKLFCKTK 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	74.1	414	1 GLYA_CAMJE	P24531 campylobact
2	41	70.7	407	1 YEAD_SCHPO	O14078 schizosacch
3	39	67.2	32	1 TX29_PHONI	P29426 phoneutria
4	39	67.2	803	1 ZN43_HUMAN	P28160 homo sapien
5	37	63.8	402	1 HEM1_METTH	O27093 methanobact
6	37	63.8	527	1 NU2W_ACACA	Q37376 acanthamoeb
7	37	63.8	328	1 YRBH_ECOLI	P45395 escherichia
8	36	62.1	704	1 Y590_METJA	Q58010 methanococc
9	35	60.3	2911	1 FBN2_HUMAN	P35556 homo sapien
10	35	60.3	2907	1 FBN2_MOUSE	O61555 mus musculu
11	35	60.3	572	1 KID1_MOUSE	Q61751 mus musculu
12	35	60.3	576	1 KID1_RAT	Q02975 rattus norv.
13	35	60.3	204	1 YM09_PARTE	P15610 paramecium
14	34.5	59.5	931	1 EMRI_MOUSE	O61549 mus musculu
15	34	58.6	420	1 GLYA_RICPR	O08370 rickettsia
16	34	58.6	424	1 SP60_DICDI	P15270 dictyosteli
17	34	58.6	118	1 YB32_METJA	Q58532 methanococc
18	34	58.6	116	1 YE80_YEAST	P40000 saccharomyc
19	34	58.6	1191	1 ZN91_HUMAN	Q05481 homo sapien
20	34	58.6	163	1 ZN92_HUMAN	Q03936 homo sapien
21	33	56.9	1130	1 ABL1_HUMAN	P00519 homo sapien
22	33	56.9	1123	1 ABL_MOUSE	P00520 mus musculu
23	33	56.9	239	1 DHSE_PORPU	P80477 porphyra pu
24	33	56.9	2871	1 FBN1_BOVIN	P98133 bos taurus
25	33	56.9	2871	1 FBN1_MOUSE	P35555 homo sapien
26	33	56.9	2871	1 FBN1_MOUSE	O61554 mus musculu
27	33	56.9	522	1 GA6S_HUMAN	P34059 homo sapien
28	33	56.9	554	1 NBLA_MOUSE	P52963 mus musculu
29	33	56.9	327	1 PGLR_BRANA	P35337 brassica na
30	33	56.9	568	1 RRP_PTIHB	P32530 human paral
31	33	56.9	568	1 RRP_PTIHC	P28054 human paral
32	33	56.9	568	1 RRP_PTIHD	P32531 human paral
33	33	56.9	568	1 RRP_PTIHE	P32532 human paral
34	33	56.9	301	1 VU51_HSV6U	P52382 herpes simp
35	33	56.9	301	1 VU51_HSV6Z	P52342 herpes simp
36	33	56.9	371	1 YKQA_CABEL	P34306 caenorhabdi
37	32.5	56.0	595	1 NASU_RAT	Q07782 rattus norv
38	32	55.2	993	1 BASO_HUMAN	O10954 homo sapien
39	32	55.2	428	1 G6NT_MOUSE	Q09324 mus musculu
40	32	55.2	691	1 NUAM_RECAM	O21241 reclinomona
41	32	55.2	889	1 RPAL_METVA	P41556 methanococc
42	32	55.2	377	1 SYW_SYNY3	P73655 synechocyst
43	32	55.2	630	1 VF12_FOWPV	P36317 fowlpox vir

ALIGNMENTS

```
RESULT 1
ID GLYA_CAMJE STANDARD; PRT; 414 AA.
AC P24531.
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DN (SHMT).
GN GLYA.
OS CAMPYLOBACTER JEJUNI.
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;
OC CAMPYLOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 43431 / TGH9011;
RX MEDLINE; 91285434.
RA CHAN V.L., BINGHAM H.L.;
RT "Complete sequence of the Campylobacter jejuni glyA gene encoding
  serine hydroxymethyltransferase."
RL GENE 101:51-58(1991).
CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
  H(2)O -> TETRAHYDROFOLATE + L-SERINE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
  HORMONES AND OTHER COMPONENTS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53816; G40534; -.
DR PIR; JQ1016; JQ1016.
DR PIR; S16468; S16468.
DR PROSITE; PS00096; SHMT; 1.
DR PFAM; PF00464; SHMT; 1.
KW TRANSFERASE; PYRIDOXAL PHOSPHATE; ONE-CARBON METABOLISM.
FT BINDING; 224 224
SQ SEQUENCE 414 AA; 45751 MW; 86662C4A CRC32;
```

Query Match 74.1%; Score 43; DB 1; Length 414;  
Best Local Similarity 87.5%; Pred. No. 1.4;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCKK 8  
|||||  
Db 77 CKKLFCKK 84

```
RESULT 2
ID YEAD_SCHPO STANDARD; PRT; 407 AA.
AC O14078.
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 45.4 KD GTP-BINDING PROTEIN C2E11.13C IN CHROMOSOME I.
```

GN SPAC2E11.13C.  
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;  
 OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;  
 OC SCHIZOSACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WOOD V.;  
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- SIMILARITY: BELONGS TO THE YCHF FAMILY OF GTP-BINDING PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL031180; E1314272; -;  
 KW HYPOTHETICAL PROTEIN; GTP-BINDING.  
 FT NP\_BIND 52 59 GTP (POTENTIAL).  
 FT NP\_BIND 106 110 GTP (POTENTIAL).  
 SQ SEQUENCE 407 AA; 45362 MW; 877F862F CRC32;

Query Match 70.7%; Score 41; DB 1; Length 407;  
 Best Local Similarity 70.0%; Pred. No. 3;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 CKKLFCKTK 10  
 || : |||||  
 Db 380 CKSVGCKTK 389

RESULT 3  
 TX29\_PHONI  
 ID TX29\_PHONI STANDARD; PRT; 32 AA.  
 AC P29426;  
 DT 01-APR-1993 (REL. 25, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
 DE NEUROTOXIN TX2-9.  
 OS PHONEUTRIA NIGRIVENTER (BRAZILIAN ARMED SPIDER).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; CHELICERATA; ARACHNIDA; ARANEAE;  
 OC ARANEOMORPHAE; ENTELEGNAE; LYCOSOIDEA; CTENIDAE; PHONEUTRIA.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=VENOM;  
 RX MEDLINE; 93011905.  
 RA CORDEIRO M.N., DINIZ C.R., VALENTIM A.C., VON EICKSTEDT V.R.D.,  
 RA GILROY J., RICHARDSON M.;  
 RT "The purification and amino acid sequences of four Tx2 neurotoxins  
 RT from the venom of the Brazilian 'armed' spider Phoneutria nigriventer  
 RT (keys).";  
 RL FEBS LETT. 310:153-156(1992).  
 CC -!- FUNCTION: CAUSES TAIL ERECTION, SCRATCHING AND A REDUCTION IN  
 CC MOBILITY AT A DOSE LEVEL OF 1.40 MG/MOUSE.  
 CC PIR; S29217; S29217.  
 DR VENOM; NEUROTOXIN.  
 SQ SEQUENCE 32 AA; 3743 MW; 3E6E1AD0 CRC32;

Query Match 67.2%; Score 39; DB 1; Length 32;  
 Best Local Similarity 77.8%; Pred. No. 0.83;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 CKKLFCKTK 9  
 | : |||||  
 Db 15 CCKKFKCKT 23

RESULT 4  
 ZN43\_HUMAN  
 ID ZN43\_HUMAN STANDARD; PRT; 803 AA.  
 AC P28150;  
 DT 01-DEC-1992 (REL. 24, CREATED)  
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE ZINC FINGER PROTEIN 43 (ZINC PROTEIN HTF6).  
 GN ZNF43.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 RC TISSUE=T-CELL;  
 RX MEDLINE; 91279444.  
 RA LOVERING R., TROWSDALE J.;  
 RT "A gene encoding 22 highly related zinc fingers is expressed in  
 RT lymphoid cell lines.";  
 RL NUCLEIC ACIDS RES. 19:2921-2927(1991).  
 RN [2]  
 RP SEQUENCE OF 38-190 FROM N.A.  
 RX MEDLINE; 91219421.  
 RA BELLEFROID E.J., PONCELET D.A., LECOQ P.J., REVELANT O.,  
 RA MARTIAL J.A.;  
 RT "The evolutionarily conserved Kruppel-associated box domain defines a  
 RT subfamily of eukaryotic multifingered proteins.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:3608-3612(1991).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- TISSUE SPECIFICITY: T AND B CELL LINES.  
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X59244; G38032; -;  
 DR EMBL; M61869; G184450; -;  
 DR PIR; S26823; S26823.  
 DR PIR; D39384; D39384.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 19.  
 DR PFAM; PF00096; zf-C2H2; 21.  
 DR HSP; P25490; 12NM.  
 KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN.  
 FT DOMAIN 1 37  
 FT KRAB BOX ("A BOX").  
 FT DOMAIN 38 69  
 FT KRAB BOX ("B BOX").  
 FT DOMAIN 168 778  
 FT ZN\_FING 168 190  
 FT ZN\_FING 196 218  
 FT ZN\_FING 224 246  
 FT ZN\_FING 252 274  
 FT ZN\_FING 280 302  
 FT ZN\_FING 308 330  
 FT ZN\_FING 336 358  
 FT ZN\_FING 364 386  
 FT ZN\_FING 392 414  
 FT ZN\_FING 420 442  
 FT ZN\_FING 448 470  
 FT ZN\_FING 476 498  
 FT ZN\_FING 504 526  
 FT ZN\_FING 532 554  
 FT ZN\_FING 560 582  
 FT ZN\_FING 588 610  
 FT ZN\_FING 616 638  
 FT ZN\_FING 644 666  
 FT ZN\_FING 672 694  
 FT ZN\_FING 700 722  
 FT ZN\_FING 722

```

FT  ZN_FING 728 750 C2H2-TYPE.
FT  ZN_FING 756 778 C2H2-TYPE.
FT  ZN_FING 181 181 P -> S (IN REF. 2).
SQ  SEQUENCE 803 AA; 93487 MW; DB052880 CRC32;

Query Match 67.2%; Score 39; DB 1; Length 803;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLFFCK 8
DB 165 KKLFFCK 171

RESULT 5
HEMI_METH STANDARD; PRT; 402 AA.
AC Q27093; 1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE GLUTAMYL-TRNA REDUCTASE (EC 1.2.1.-) (GLUTR).
GN HEMA OR MTH1012.
OS METHANOBACTERIUM THERMOAUTOTROPHICUM.
OC ARCHAEA; EURYARCHAEOTA; METHANOBACTERIALES; METHANOBACTERIACEAE;
OC METHANOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE; 98037514.
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA ALDRIDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA HARRISON D., HOANG L., KEAGLE P., LUM M.W., POTHIER B., QIU D.,
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. BACTERIOL. 179:7135-7155(1997).
CC -1- CATALYTIC ACTIVITY: GLUTAMYL-TRNA(GLU) + NADPH -> GLUTAMATE-1-
CC -1- SEMIALDEHYDE + NADP(+) + TRNA(GLU).
CC -1- PATHWAY: FIRST STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AE000874; G2622114; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW PORPHYRIN BIOSYNTHESIS; OXIDOREDUCTASE; NADP.
SQ SEQUENCE 402 AA; 44940 MW; 56A78E87 CRC32;

Query Match 63.8%; Score 37; DB 1; Length 402;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFFCK 8
DB 392 CAKLFFCK 399

RESULT 6
NU2M_ACACA STANDARD; PRT; 527 AA.
ID NU2M_ACACA

```

```

AC Q37376;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
GN ND2 OR NAD2.
OS ACANTHAMOEBA CASTELLANII (AMOEBA).
OC MITOCHONDRION.
OC EUKARYOTA; ACANTHAMOEBAE; ACANTHAMOEBA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 30010 / NEFF.
RA BURGER G., PLANTE I., LONGERAN K.M., GRAY M.W.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; U12386; G562039; 1.
DR PFAM; PF00361; oxidored_g1; 1.
KW OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION; TRANSMEMBRANE.
SQ SEQUENCE 527 AA; 61407 MW; D7FC10C8 CRC32;

Query Match 63.8%; Score 37; DB 1; Length 527;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKLFFCKT 9
DB 13 KKLFFCKT 20

RESULT 7
YRBH_ECOLI STANDARD; PRT; 328 AA.
ID YRBH_ECOLI
AC P45395;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL 35.2 KD PROTEIN IN MURA-RPON INTERGENIC REGION (O328).
GN YRBH.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RT SCIENCE 277:1453-1474(1997).
CC -1- SIMILARITY: STRONG, TO H. INFLUENZAE HI1678.
CC -1- SIMILARITY: TO E. COLI GUT0 AND KPSF.
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL; U18997; G606135; -.  
 DR EMBL; AE000399; G1789588; -.  
 DR ECOGENE; EG12803; YRBH.  
 DR PFAM; PF00571; CBS; 2.  
 KW HYPOTHETICAL PROTEIN; ATP-BINDING; REPEAT; CBS DOMAIN.  
 FT NP\_BIND 56 61 ATP (POTENTIAL).  
 FT DOMAIN 208 263 CBS 1.  
 FT DOMAIN 275 327 CBS 1.  
 SQ SEQUENCE 328 AA; 35196 MW; E7507FCA CRC32;

Query Match 63.8%; Score 37; DB 1; Length 328;  
 Best Local Similarity 60.0%; Pred. No. 12;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKKLFCKTK 10  
 Db 42 CEKMFCKGK 51

RESULT 8  
 Y590\_METJA STANDARD; PRT; 704 AA.  
 AC Q58010;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN M40590.  
 GN M40590.  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 CC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUITON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEORHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii".  
 RL SCIENCE 273:1058-1073(1996).  
 CC -!- SIMILARITY: SOME, IN THE N-TERMINAL, TO SUCCINYL-COA LIGASES.  
 CC -!- SIMILARITY: STRONG, TO E.COLI YFIQ.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; U67507; G1591298; -.  
 DR TIGR; MJ0590; -.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 704 AA; 78172 MW; 148F4430 CRC32;

Query Match 62.1%; Score 36; DB 1; Length 704;  
 Best Local Similarity 70.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CKKLFCKTK 10  
 Db 521 CKKLGCKYWK 530

RESULT 9  
 FBN2\_HUMAN STANDARD; PRT; 2911 AA.  
 AC P35556;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE FIBRILLIN 2 PRECURSOR.  
 GN FBN2.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94165150.  
 RA ZHANG H., APFELROTH S.D., HU W., DAVIS E.C., SANGUINETI C.,  
 RA BONADIO J., MECHAM R.P., RAMIREZ F.;  
 RT "Structure and expression of fibrillin-2, a novel microfibrillar  
 component preferentially located in elastic matrices.";  
 RL J. CELL BIOL. 124:855-863(1994).  
 RN [2]  
 RP SEQUENCE OF 752-1505 FROM N.A.  
 RX MEDLINE; 91304567.  
 RA LEE B., GODFREY M., VITALE E., HORI H., MATTEI M.-G., SARFARAZI M.,  
 RA TSIPOURAS P., RAMIREZ F., HOLLISTER D.;  
 RT "Linkage of Marfan syndrome and a phenotypically related disorder to  
 two different fibrillin genes.";  
 RL NATURE 352:330-334(1991).  
 RN [3]  
 RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.  
 RX MEDLINE; 96083599.  
 RA PUTNAM E.A., ZHANG H., RAMIREZ F., MILEWICZ D.M.;  
 RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,  
 congenital contractual arachnoidecty.";  
 RL NAT. GENET. 11:456-458(1995).  
 CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
 THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE  
 THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.  
 CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL  
 ARACHNOECTY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS  
 PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE  
 AORTA AND THE EYES.  
 CC -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING  
 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; U03272; G437972; -.  
 DR EMBL; X62009; -. NOT\_ANNOTATED\_CDS.  
 DR PIR; S17063; S17063.  
 DR PIR; S31101; S31101.  
 DR MIN; I21050; -.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 43.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 37.  
 DR PROSITE; PS01187; EGF\_CA; 43.  
 DR PFAM; PF00008; EGF; 42.  
 DR PFAM; PF00683; TB; 9.  
 DR HSP; P35555; IEMO.  
 KW EXTRACELLULAR MATRIX; CALCIUM-BINDING; GLYCOPROTEIN; EGF-LIKE DOMAIN;  
 KW REPEAT; SIGNAL; MULTIGENE FAMILY; DISEASE MUTATION; POLYMORPHISM.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 2911 FIBRILLIN 2.  
 FT DOMAIN 111 142 EGF-LIKE 1, NON-CALCIUM BINDING.  
 FT DOMAIN 145 176 EGF-LIKE 2, NON-CALCIUM BINDING.  
 FT DOMAIN 176 207 EGF-LIKE 3, NON-CALCIUM BINDING.

FT DOMAIN	275	EGF-LIKE 4, CALCIUM-BINDING.	FT DISULFID	585	599	BY SIMILARITY.
FT DOMAIN	317	EGF-LIKE 5, CALCIUM-BINDING.	FT DISULFID	601	614	BY SIMILARITY.
FT REPEAT	359	TGFBP 1.	FT DISULFID	620	631	BY SIMILARITY.
FT DOMAIN	493	EGF-LIKE 6, NON-CALCIUM BINDING.	FT DISULFID	626	640	BY SIMILARITY.
FT DOMAIN	534	EGF-LIKE 7, CALCIUM-BINDING.	FT DISULFID	642	655	BY SIMILARITY.
FT DOMAIN	574	EGF-LIKE 8, CALCIUM-BINDING.	FT DISULFID	661	672	BY SIMILARITY.
FT DOMAIN	616	EGF-LIKE 9, CALCIUM-BINDING.	FT DISULFID	667	681	BY SIMILARITY.
FT DOMAIN	657	EGF-LIKE 10, CALCIUM-BINDING.	FT DISULFID	683	696	BY SIMILARITY.
FT REPEAT	698	TGFBP 2.	FT DISULFID	771	782	BY SIMILARITY.
FT DOMAIN	767	EGF-LIKE 11, CALCIUM-BINDING.	FT DISULFID	778	792	BY SIMILARITY.
FT DOMAIN	809	EGF-LIKE 12, CALCIUM-BINDING.	FT DISULFID	794	807	BY SIMILARITY.
FT DOMAIN	851	EGF-LIKE 13, CALCIUM-BINDING.	FT DISULFID	813	825	BY SIMILARITY.
FT DOMAIN	954	EGF-LIKE 14, CALCIUM-BINDING.	FT DISULFID	820	834	BY SIMILARITY.
FT REPEAT	996	TGFBP 3.	FT DISULFID	836	849	BY SIMILARITY.
FT DOMAIN	1072	EGF-LIKE 15, CALCIUM-BINDING.	FT DISULFID	855	865	BY SIMILARITY.
FT DOMAIN	1114	EGF-LIKE 16, CALCIUM-BINDING.	FT DISULFID	860	874	BY SIMILARITY.
FT DOMAIN	1157	EGF-LIKE 17, CALCIUM-BINDING.	FT DISULFID	876	889	BY SIMILARITY.
FT DOMAIN	1199	EGF-LIKE 18, CALCIUM-BINDING.	FT DISULFID	876	890	BY SIMILARITY.
FT DOMAIN	1240	EGF-LIKE 19, CALCIUM-BINDING.	FT DISULFID	958	970	BY SIMILARITY.
FT DOMAIN	1241	EGF-LIKE 20, CALCIUM-BINDING.	FT DISULFID	965	979	BY SIMILARITY.
FT DOMAIN	1282	EGF-LIKE 21, CALCIUM-BINDING.	FT DISULFID	981	994	BY SIMILARITY.
FT DOMAIN	1324	EGF-LIKE 22, CALCIUM-BINDING.	FT DISULFID	1076	1088	BY SIMILARITY.
FT DOMAIN	1366	EGF-LIKE 23, CALCIUM-BINDING.	FT DISULFID	1083	1097	BY SIMILARITY.
FT DOMAIN	1407	EGF-LIKE 24, CALCIUM-BINDING.	FT DISULFID	1099	1112	BY SIMILARITY.
FT DOMAIN	1448	EGF-LIKE 25, CALCIUM-BINDING.	FT DISULFID	1118	1130	BY SIMILARITY.
FT DOMAIN	1490	EGF-LIKE 26, CALCIUM-BINDING.	FT DISULFID	1125	1139	BY SIMILARITY.
FT REPEAT	1531	TGFBP 4.	FT DISULFID	1141	1155	BY SIMILARITY.
FT REPEAT	1572	EGF-LIKE 27, CALCIUM-BINDING.	FT DISULFID	1161	1173	BY SIMILARITY.
FT DOMAIN	1649	EGF-LIKE 28, CALCIUM-BINDING.	FT DISULFID	1168	1182	BY SIMILARITY.
FT DOMAIN	1691	EGF-LIKE 29, CALCIUM-BINDING.	FT DISULFID	1184	1197	BY SIMILARITY.
FT REPEAT	1733	TGFBP 5.	FT DISULFID	1203	1215	BY SIMILARITY.
FT DOMAIN	1807	EGF-LIKE 30, CALCIUM-BINDING.	FT DISULFID	1210	1224	BY SIMILARITY.
FT DOMAIN	1849	EGF-LIKE 31, CALCIUM-BINDING.	FT DISULFID	1226	1239	BY SIMILARITY.
FT DOMAIN	1891	EGF-LIKE 32, CALCIUM-BINDING.	FT DISULFID	1245	1256	BY SIMILARITY.
FT DOMAIN	1933	EGF-LIKE 33, CALCIUM-BINDING.	FT DISULFID	1252	1265	BY SIMILARITY.
FT DOMAIN	1972	EGF-LIKE 34, CALCIUM-BINDING.	FT DISULFID	1267	1280	BY SIMILARITY.
FT DOMAIN	2015	EGF-LIKE 35, CALCIUM-BINDING.	FT DISULFID	1286	1298	BY SIMILARITY.
FT DOMAIN	2055	EGF-LIKE 36, CALCIUM-BINDING.	FT DISULFID	1293	1307	BY SIMILARITY.
FT REPEAT	2097	TGFBP 6.	FT DISULFID	1309	1322	BY SIMILARITY.
FT DOMAIN	2170	EGF-LIKE 37, CALCIUM-BINDING.	FT DISULFID	1328	1340	BY SIMILARITY.
FT DOMAIN	2212	EGF-LIKE 38, CALCIUM-BINDING.	FT DISULFID	1335	1349	BY SIMILARITY.
FT DOMAIN	2252	EGF-LIKE 39, CALCIUM-BINDING.	FT DISULFID	1351	1364	BY SIMILARITY.
FT DOMAIN	2293	EGF-LIKE 40, CALCIUM-BINDING.	FT DISULFID	1370	1383	BY SIMILARITY.
FT DOMAIN	2337	EGF-LIKE 41, CALCIUM-BINDING.	FT DISULFID	1377	1392	BY SIMILARITY.
FT REPEAT	2379	TGFBP 7.	FT DISULFID	1394	1405	BY SIMILARITY.
FT DOMAIN	2448	EGF-LIKE 42, CALCIUM-BINDING.	FT DISULFID	1411	1424	BY SIMILARITY.
FT DOMAIN	2489	EGF-LIKE 43, CALCIUM-BINDING.	FT DISULFID	1418	1433	BY SIMILARITY.
FT DOMAIN	2530	EGF-LIKE 44, CALCIUM-BINDING.	FT DISULFID	1435	1446	BY SIMILARITY.
FT DOMAIN	2531	EGF-LIKE 45, CALCIUM-BINDING.	FT DISULFID	1452	1464	BY SIMILARITY.
FT DOMAIN	2570	EGF-LIKE 46, CALCIUM-BINDING.	FT DISULFID	1459	1473	BY SIMILARITY.
FT DOMAIN	2613	EGF-LIKE 47, CALCIUM-BINDING.	FT DISULFID	1475	1488	BY SIMILARITY.
FT DOMAIN	2653	BY SIMILARITY.	FT DISULFID	1494	1505	BY SIMILARITY.
FT DOMAIN	2694	BY SIMILARITY.	FT DISULFID	1500	1514	BY SIMILARITY.
FT DISULFID	115	BY SIMILARITY.	FT DISULFID	1516	1529	BY SIMILARITY.
FT DISULFID	119	BY SIMILARITY.	FT DISULFID	1535	1546	BY SIMILARITY.
FT DISULFID	132	BY SIMILARITY.	FT DISULFID	1541	1555	BY SIMILARITY.
FT DISULFID	149	BY SIMILARITY.	FT DISULFID	1557	1570	BY SIMILARITY.
FT DISULFID	153	BY SIMILARITY.				
FT DISULFID	166	BY SIMILARITY.				
FT DISULFID	180	BY SIMILARITY.				
FT DISULFID	184	BY SIMILARITY.				
FT DISULFID	195	BY SIMILARITY.				
FT DISULFID	197	BY SIMILARITY.				
FT DISULFID	206	BY SIMILARITY.				
FT DISULFID	279	BY SIMILARITY.				
FT DISULFID	286	BY SIMILARITY.				
FT DISULFID	302	BY SIMILARITY.				
FT DISULFID	321	BY SIMILARITY.				
FT DISULFID	328	BY SIMILARITY.				
FT DISULFID	344	BY SIMILARITY.				
FT DISULFID	497	BY SIMILARITY.				
FT DISULFID	504	BY SIMILARITY.				
FT DISULFID	520	BY SIMILARITY.				
FT DISULFID	538	BY SIMILARITY.				
FT DISULFID	543	BY SIMILARITY.				
FT DISULFID	559	BY SIMILARITY.				
FT DISULFID	578	BY SIMILARITY.				

Query Match 60.38; Score 35; DB 1; Length 2911;  
Best Local Similarity 60.08; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKLFLCKTK 10  
|||:|:  
Db 2529 CKDLDECQTK 2538

RESULT 10  
FBN2\_MOUSE  
ID FBN2\_MOUSE STANDARD; PRT: 2907 AA.  
AC Q61555; Q63957;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)





FT DISULFID 1093 BY SIMILARITY. 1106  
 FT DISULFID 1112 BY SIMILARITY. 1124  
 FT DISULFID 1119 BY SIMILARITY. 1133  
 FT DISULFID 1135 BY SIMILARITY. 1149  
 FT DISULFID 1155 BY SIMILARITY. 1167  
 FT DISULFID 1162 BY SIMILARITY. 1176  
 FT DISULFID 1178 BY SIMILARITY. 1191  
 FT DISULFID 1197 BY SIMILARITY. 1209  
 FT DISULFID 1204 BY SIMILARITY. 1218  
 FT DISULFID 1220 BY SIMILARITY. 1233  
 FT DISULFID 1239 BY SIMILARITY. 1250  
 FT DISULFID 1246 BY SIMILARITY. 1259  
 FT DISULFID 1261 BY SIMILARITY. 1274  
 FT DISULFID 1280 BY SIMILARITY. 1292  
 FT DISULFID 1287 BY SIMILARITY. 1301  
 FT DISULFID 1303 BY SIMILARITY. 1316  
 FT DISULFID 1322 BY SIMILARITY. 1334  
 FT DISULFID 1329 BY SIMILARITY. 1343  
 FT DISULFID 1345 BY SIMILARITY. 1358  
 FT DISULFID 1364 BY SIMILARITY. 1377  
 FT DISULFID 1371 BY SIMILARITY. 1386  
 FT DISULFID 1388 BY SIMILARITY. 1399  
 FT DISULFID 1405 BY SIMILARITY. 1418  
 FT DISULFID 1412 BY SIMILARITY. 1427  
 FT DISULFID 1429 BY SIMILARITY. 1440  
 FT DISULFID 1446 BY SIMILARITY. 1458  
 FT DISULFID 1453 BY SIMILARITY. 1467  
 FT DISULFID 1468 BY SIMILARITY. 1482  
 FT DISULFID 1488 BY SIMILARITY. 1499  
 FT DISULFID 1494 BY SIMILARITY. 1508  
 FT DISULFID 1510 BY SIMILARITY. 1523  
 FT DISULFID 1529 BY SIMILARITY. 1540  
 FT DISULFID 1535 BY SIMILARITY. 1549  
 FT DISULFID 1551 BY SIMILARITY. 1564  
 FT DISULFID 1647 BY SIMILARITY. 1659  
 FT DISULFID 1654 BY SIMILARITY. 1668  
 FT DISULFID 1670 BY SIMILARITY. 1683  
 FT DISULFID 1689 BY SIMILARITY. 1701  
 FT DISULFID 1696 BY SIMILARITY. 1710  
 FT DISULFID 1712 BY SIMILARITY. 1725  
 FT DISULFID 1805 BY SIMILARITY. 1817  
 FT DISULFID 1812 BY SIMILARITY. 1826  
 FT DISULFID 1828 BY SIMILARITY. 1841  
 FT DISULFID 1847 BY SIMILARITY. 1860  
 FT DISULFID 1854 BY SIMILARITY. 1869  
 FT DISULFID 1871 BY SIMILARITY. 1883  
 FT DISULFID 1889 BY SIMILARITY. 1901  
 FT DISULFID 1896 BY SIMILARITY. 1910  
 FT DISULFID 1912 BY SIMILARITY. 1925

Query Match 60.3%; Score 35; DB 1; Length 2907;  
 Best Local Similarity 60.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CKLFCKTK 10  
 ||| :||  
 Db 2523 CKLDECOTK 2532

RESULT 11  
 KID1\_MOUSE  
 ID KID1\_MOUSE STANDARD; PRT; 572 AA.  
 AC Q61751;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17).  
 GN TCF17 OR KID1.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE-LENS;  
 RA BRADY J.P., DUNCAN M.K., WAMROUSEK E.F., PIATIGORSKY J.;  
 RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- FUNCTION: IT MAY PLAY A ROLE IN RENAL DEVELOPMENT AND MAY  
 CC ALSO BE INVOLVED IN THE REPAIR OF THE KIDNEY AFTER ISCHEMIA-  
 CC REPERFUSION OR FOLIC ACID ADMINISTRATION.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 1 KRAB BOX.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC  
 CC EMBL; L77247; G1256362; .  
 CC MGD; MGI:103172; TCF17.  
 CC PROSITE: PS00028; ZINC\_FINGER\_C2H2; 13.  
 CC PFAM; PF00096; zf-C2H2; 13.  
 CC HSP; P08047; ISP2.  
 CC ZINC-FINGER; TRANSCRIPTION REGULATION; DNA-BINDING; ACTIVATOR;  
 CC NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN; METAL-BINDING.  
 CC DOMAIN 12 53 KRAB BOX ("A BOX").  
 CC DOMAIN 54 85 KRAB BOX ("B BOX").  
 CC DOMAIN 184 567 ZINC-FINGERS.  
 CC ZN\_FING 184 206 C2H2-TYPE.  
 CC ZN\_FING 212 234 C2H2-TYPE.  
 CC ZN\_FING 240 262 C2H2-TYPE.  
 CC ZN\_FING 268 290 C2H2-TYPE.  
 CC ZN\_FING 321 343 C2H2-TYPE.  
 CC ZN\_FING 349 371 C2H2-TYPE.  
 CC ZN\_FING 377 399 C2H2-TYPE.  
 CC ZN\_FING 405 427 C2H2-TYPE.  
 CC ZN\_FING 431 455 C2H2-TYPE.  
 CC ZN\_FING 461 483 C2H2-TYPE.  
 CC ZN\_FING 489 511 C2H2-TYPE.  
 CC ZN\_FING 517 539 C2H2-TYPE.  
 CC ZN\_FING 545 567 C2H2-TYPE.  
 CC SEQUENCE 572 AA; 65720 MW; 4F7EBBDC CRC32;

Query Match 60.3%; Score 35; DB 1; Length 572;  
 Best Local Similarity 85.7%; Pred. No. 41;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KLFCKK 8  
 :|||||  
 Db 209 KLFCKK 215

RESULT 12  
 KID1\_RAT  
 ID KID1\_RAT STANDARD; PRT; 576 AA.  
 AC Q02975;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17).  
 GN TCF17 OR KID1  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-KIDNEY;  
 RX MEDLINE; 93180842.  
 RA WITZGALL R., O'LEARY E., GESSNER R., OUELLETTE A.J., BONVENTRE J.V.;  
 RT "Kid-1, a putative renal transcription factor: regulation during  
 RT ontogeny and in response to ischemia and toxic injury.";

RL MOL. CELL. BIOL. 13:1933-1942(1993).  
 RP [2]  
 RN SEQUENCE OF 1-110 FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE; 94292199.  
 RA WITZGALL R., VOLK R., YEUNG R.S., BONVENTRE J.V.;  
 RT "Genomic structure and chromosomal location of the rat gene encoding  
 the zinc finger transcription factor Kid-1.";  
 RL GENOMICS 20:203-209(1994).  
 CC -!- FUNCTION: IT MAY PLAY A ROLE IN RENAL DEVELOPMENT AND MAY  
 ALSO BE INVOLVED IN THE REPAIR OF THE KIDNEY AFTER ISCHEMIA-  
 REPERFUSION OR FOLIC ACID ADMINISTRATION.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN THE KIDNEY.  
 CC -!- DEVELOPMENTAL STAGE: PROTEIN LEVELS ACCUMULATE WITH AGE IN  
 POSTNATAL RENAL DEVELOPMENT.  
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-  
 FINGER PROTEINS.  
 CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M96548; G205067; .  
 DR EMBL; S73488; G688395; .  
 DR EMBL; S73484; G688395; JOINED.  
 DR EMBL; S73485; G688395; JOINED.  
 DR EMBL; S73486; G688395; JOINED.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 13.  
 DR PFAM; PF00096; 1zf-C2H2; 13.  
 DR HSP; P08047; 1SP2.  
 DR TRANSFAC; T02326; .  
 KW ZINC-FINGER; TRANSCRIPTION REGULATION; DNA-BINDING; ACTIVATOR;  
 KW NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN; METAL-BINDING.  
 FT DOMAIN 12 54  
 FT DOMAIN 55 86  
 FT DOMAIN 96 111  
 FT DOMAIN 170 189  
 FT DOMAIN 190 552  
 FT ZN\_FING 190 210  
 FT ZN\_FING 218 238  
 FT ZN\_FING 246 266  
 FT ZN\_FING 274 294  
 FT ZN\_FING 308 328  
 FT ZN\_FING 336 356  
 FT ZN\_FING 364 384  
 FT ZN\_FING 392 412  
 FT ZN\_FING 420 440  
 FT ZN\_FING 448 468  
 FT ZN\_FING 476 496  
 FT ZN\_FING 504 524  
 FT ZN\_FING 532 552  
 SQ SEQUENCE 576 AA; 66253 MW; 6F583606 CRC32;

Query Match 60.3%; Score 35; DB 1; Length 576;  
 Best Local Similarity 85.7%; Pred. No. 42;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KKLKCK 8  
 Db 213 EKLKCK 219

RESULT 13  
 YMO9\_PARTE  
 ID YMO9\_PARTE  
 AC P15610; PRT; 204 AA.

DT 01-APR-1990 (REL. 14, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DE HYPOTHETICAL 22.9 KD PROTEIN (ORF9).  
 OS PARAMECIUM TETRAURELIA.  
 OG MITOCHONDRION.  
 CC EUKARYOTA; ALVEOLATA; CILIOPHORA; NASSOPHOREA; PENICULIDA; PARAMECIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-STOCK 51;  
 RX MEDLINE; 90174913.  
 RA PRITCHARD A.E., SEILHAMER J.J., MAHALINGAM R., SABLE C.L.,  
 RA VENUTI S.E., CUMINGS D.J.;  
 RT "Nucleotide sequence of the mitochondrial genome of Paramecium.";  
 RL NUCLEIC ACIDS RES. 18:173-180(1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X15917; G578758; .  
 DR PIR; S07741; S07741.  
 KW HYPOTHETICAL PROTEIN; MITOCHONDRION.  
 SQ SEQUENCE 204 AA; 22919 MW; 27F9BFCF CRC32;

Query Match 60.3%; Score 35; DB 1; Length 204;  
 Best Local Similarity 60.0%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CKLFLKCKTK 10  
 Db 114 CKKFFKLKR 123

RESULT 14

EMR1\_MOUSE  
 ID EMR1\_MOUSE STANDARD; PRT; 931 AA.  
 AC O61549;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE CELL SURFACE GLYCOPROTEIN EMR1 PRECURSOR (EMR1 HORMONE RECEPTOR)  
 DE (CELL SURFACE GLYCOPROTEIN F4/80).  
 GN EMR1 OR GPF480.  
 OS MUS MUSCULUS (MOUSE).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE=PERITONEAL CAVITY;  
 RX MEDLINE; 96132946.  
 RA MCKNIGHT A.J., MACFARLANE A.J., DRI P., TURLEY L., WILLIS A.C.,  
 RA GORDON S.;  
 RT "Molecular cloning of F4/80, a murine macrophage-restricted cell  
 surface glycoprotein with homology to the G-protein-linked  
 transmembrane 7 hormone receptor family.";  
 RL J. BIOL. CHEM. 271:486-489(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97312684.  
 RA LIN H.H., STUBBS L.J., MUCENSKI M.L.;  
 RT "Identification and characterization of a seven transmembrane hormone  
 receptor using differential display.";  
 RL GENOMICS 41:301-308(1997).  
 CC -!- FUNCTION: PROBABLY INVOLVED IN CELL ADHESION WITHIN TISSUES  
 AND RECEPTOR SIGNALING.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: IN MACROPHAGES; BUT ABSENT FROM THOSE WHICH

CC ARE LOCALIZED WITHIN T-CELL AREAS OF LYMPH NODES AND SPLEEN.  
 CC LOW LEVEL OF EXPRESSION ON BLOOD MONOCYTES.  
 CC -1- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X93328; E214264; -;  
 CC DR EMBL; U66888; G2078504; -;  
 CC DR GCRDB; GCR1309; -;  
 CC DR MGD; MGI-105054; GPF480.  
 CC DR PROSITE; PS00649; G\_PROTEIN\_RECEPT\_F2\_1; FALSE\_NEG.  
 CC DR PROSITE; PS00650; G\_PROTEIN\_RECEPT\_F2\_2; 1.  
 CC DR PROSITE; PS00010; ASX\_HYDROXYL; 6.  
 CC DR PROSITE; PS01186; EGF\_2; 1.  
 CC DR PROSITE; PS01187; EGF\_CA; 5.  
 CC DR PFAM; PF00008; EGF; 1.  
 CC DR HSP; P07204; IFGD.  
 CC KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; RECEPTOR; GLYCOPROTEIN;  
 CC EGF-LIKE DOMAIN; REPEAT; SIGNAL.  
 CC SIGNAL 1 27 POTENTIAL.  
 CC FT CHAIN 28 931 CELL SURFACE GLYCOPROTEIN F4/80.  
 CC FT DOMAIN 28 644 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSEM 645 672 POTENTIAL.  
 CC FT DOMAIN 673 679 CYTOPLASMIC (POTENTIAL).  
 CC FT TRANSEM 680 701 POTENTIAL.  
 CC FT DOMAIN 702 711 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSEM 712 735 POTENTIAL.  
 CC FT DOMAIN 736 754 CYTOPLASMIC (POTENTIAL).  
 CC FT TRANSEM 755 776 POTENTIAL.  
 CC FT DOMAIN 777 792 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSEM 793 821 POTENTIAL.  
 CC FT DOMAIN 822 839 CYTOPLASMIC (POTENTIAL).  
 CC FT TRANSEM 840 859 POTENTIAL.  
 CC FT DOMAIN 860 874 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSEM 875 897 POTENTIAL.  
 CC FT DOMAIN 898 931 CYTOPLASMIC (POTENTIAL).  
 CC FT SITE 32 80 EGF-LIKE 1.  
 CC FT DOMAIN 81 132 EGF-LIKE 2.  
 CC FT DOMAIN 133 172 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 CC FT DOMAIN 173 221 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 CC FT DOMAIN 222 271 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 CC FT DOMAIN 272 318 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
 CC FT DOMAIN 319 367 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).  
 CC FT SITE 506 508 CELL ATTACHMENT SITE (POTENTIAL).  
 CC FT DISULFID 36 48 BY SIMILARITY.  
 CC FT DISULFID 42 57 BY SIMILARITY.  
 CC FT DISULFID 59 79 BY SIMILARITY.  
 CC FT DISULFID 85 98 BY SIMILARITY.  
 CC FT DISULFID 92 107 BY SIMILARITY.  
 CC FT DISULFID 109 131 BY SIMILARITY.  
 CC FT DISULFID 137 149 BY SIMILARITY.  
 CC FT DISULFID 143 158 BY SIMILARITY.  
 CC FT DISULFID 160 171 BY SIMILARITY.  
 CC FT DISULFID 177 189 BY SIMILARITY.  
 CC FT DISULFID 183 198 BY SIMILARITY.  
 CC FT DISULFID 200 220 BY SIMILARITY.  
 CC FT DISULFID 226 239 BY SIMILARITY.  
 CC FT DISULFID 233 248 BY SIMILARITY.  
 CC FT DISULFID 250 270 BY SIMILARITY.  
 CC FT DISULFID 276 286 BY SIMILARITY.  
 CC FT DISULFID 280 295 BY SIMILARITY.  
 CC FT DISULFID 297 317 BY SIMILARITY.  
 CC FT DISULFID 323 336 BY SIMILARITY.  
 CC FT DISULFID 330 345 BY SIMILARITY.  
 CC FT DISULFID 347 366 BY SIMILARITY.  
 CC FT CARBOHYD 148 148 POTENTIAL.

FT CARBOHYD 167 167 POTENTIAL.  
 FT CARBOHYD 229 229 POTENTIAL.  
 FT CARBOHYD 269 269 POTENTIAL.  
 FT CARBOHYD 283 283 POTENTIAL.  
 FT CARBOHYD 405 405 POTENTIAL.  
 FT CARBOHYD 417 417 POTENTIAL.  
 FT CARBOHYD 474 474 POTENTIAL.  
 FT CARBOHYD 498 498 POTENTIAL.  
 FT CARBOHYD 706 706 POTENTIAL.  
 SQ SEQUENCE 931 AA; 102129 MW; 35DB95D7 CRC32;  
 Query Match 59.5%; Score 34.5; DB 1; Length 931;  
 Best Local Similarity 77.8%; Pred. No. 75;  
 Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 CKK-LFKCK 8  
 DB 366 CKRLEKCK 374  
 ||: |||||  
 ||: |||||  
 RESULT 15  
 GLYA\_RICPR STANDARD; PRT; 420 AA.  
 ID GLYA\_RICPR AC 008370;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)  
 DE (SHMT).  
 OS RICKETTSIA PROWAZEKII.  
 GN RICKETTSIA PROWAZEKII.  
 OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RICKETTSIALES;  
 OC RICKETTSIAE; RICKETTSIAE; RICKETTSIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-MADRID E;  
 RX MEDLINE; 97419517.  
 RA ANDERSSON J.O., ANDERSSON S.G.E.;  
 RT "Genomic rearrangements during evolution of the obligate  
 RT intracellular parasite Rickettsia prowazekii as inferred from an  
 RT analysis of 52015 bp nucleotide sequence.";  
 RL MICROBIOLOGY 143:2783-2795(1997).  
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +  
 CC H(2)O -> TETRAHYDROFOLATE + L-SERINE.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,  
 CC HORMONES AND OTHER COMPONENTS.  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Y11778; E307841; -;  
 CC DR PROSITE; PS00096; SHMT; 1.  
 CC DR PFAM; PF00464; SHMT; 1.  
 CC KW TRANSFERASE; PYRIDOXAL PHOSPHATE; ONE-CARBON METABOLISM.  
 FT BINDING 230 230 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 SQ SEQUENCE 420 AA; 46381 MW; 4C8ACBEE CRC32;  
 Query Match 58.6%; Score 34; DB 1; Length 420;  
 Best Local Similarity 85.7%; Pred. No. 48;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KKLFCK 8  
    ||| ||  
Db 83 KLENCK 89

Search completed: September 7, 1999, 23:59:15  
Job time: 527 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run On: September 7, 1999, 22:47:44 ; Search time 160.2 Seconds  
(without alignments)  
3.842 Million cell updates/sec

Title: US-09-124-280A-33

Perfect score: 58

Sequence: 1 CKKLFCKTK 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- SPTREMBL\_10:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertibrate:\*
  - 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	74.1	1914	4	Q13129 homo sapien
2	38	65.5	241	2	O67793 aquifex ao
3	38	65.5	557	4	O43340 homo sapien
4	37	63.8	318	2	Q56914 yersinia en
5	37	63.8	1077	3	O74853 schizosacch
6	37	63.8	1051	5	Q26055 paracentrot
7	37	63.8	774	5	O18387 drosophila
8	37	63.8	3198	5	Q26639 strongyloce
9	37	63.8	999	12	Q86519 rice yellow
10	37	63.8	999	12	Q86525 rice yellow
11	36	62.1	200	5	Q94993 drosophila
12	35	60.3	1087	4	O43279 homo sapien
13	35	60.3	3703	4	Q15911 homo sapien
14	35	60.3	605	4	O60765 homo sapien
15	35	60.3	2783	4	Q13719 homo sapien
16	35	60.3	1087	4	Q99639 homo sapien
17	35	60.3	1087	4	O00267 homo sapien
18	35	60.3	299	5	Q23584 caenorhabdi
19	35	60.3	608	5	Q23898 dictyostell
20	35	60.3	608	5	O96849 dictyostell
21	35	60.3	146	10	O22087 petunia hyb
22	35	60.3	163	10	O22088 petunia hyb
23	35	60.3	311	10	Q920G6 arabidopsis
24	35	60.3	402	10	Q920G5 arabidopsis
25	35	60.3	3726	11	Q61329 mus musculu
26	34.5	59.5	297	11	O08743 mus musculu
27	34.5	59.5	304	11	O08744 mus musculu
28	34.5	59.5	255	11	O08745 mus musculu
29	34	58.6	770	3	Q06833 saccharomyc

ALIGNMENTS

RESULT 1

Q13129 PRELIMINARY; PRT; 1914 AA.  
ID Q13129  
AC Q13129;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE ZN-15 RELATED ZINC FINGER PROTEIN (RLF).  
GN RLF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 9613723.  
RA MAKELA T.P., HELLISTEN E., VESA J., HIRVONEN H., PALOTIE A.,  
RA PELTONEN L., ALITALO K., ALITALO K.;  
RT "The rearranged l-myc fusion gene (RLF) encodes a Zn-15 related zinc  
RT finger protein.";  
RL Oncogene 11:2699-2704(1995).  
DR EMBL; U22377; AAC50396.1; -.  
DR PFAM; PF00096; zf-C2H2; 14.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 14.  
KW Zinc-finger; Metal-binding; DNA-binding.  
SQ SEQUENCE 1914 AA; 217898 MW; 05EBE74D CRC32;

Query Match 74.1%; Score 43; DB 4; Length 1914;  
Best Local Similarity 70.0%; Pred. No. 4.7;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10

|||||  
Db 1358 CKKIFACKYK 1367

RESULT 2

O67793 PRELIMINARY; PRT; 241 AA.  
ID O67793  
AC O67793;  
DT 01-AUG-1998 (TREMREL. 07, Created)  
DT 01-AUG-1998 (TREMREL. 07, Last sequence update)  
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)  
DE HYPOTHETICAL 27.5 KD PROTEIN.  
GN AQ\_1986.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF5;  
RX MEDLINE; 98196666.  
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,

RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RL aeolicus.";  
 RN Nature 392:353-358(1998).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 DL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE000766; AAC07762.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 241 AA; 27514 MW; A94A054B CRC32;

Query Match 65.5%; Score 38; DB 2; Length 241;  
 Best Local Similarity 75.0%; Pred. No. 7.8;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCK 8  
 |||||  
 DB 162 CEKLLCK 169

RESULT 3  
 ID Q43340 PRELIMINARY; PRT; 557 AA.  
 AC Q43340;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE R28830\_2  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LAMERDIN J.E., MCCREADY P.M., ADAMSON A.W., BURKHART-SCHULTZ K.,  
 RA GARCIA E., KYLE A., RAMIREZ M., STILWAGEN S., GARNES J., DANGANAN L.,  
 RA BRUCE R., QUAN G., MONTGOMERY M., OW D., KOBAYASHI A., OLSEN A.O.,  
 RA CARRANO A.V.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC003682; AAC24609.1; -;  
 DR PFAM; PF00096; zf-C2H2; 13.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 13.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 557 AA; 63861 MW; ED322DFA CRC32;

Query Match 65.5%; Score 38; DB 4; Length 557;  
 Best Local Similarity 66.7%; Pred. No. 15;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCK 9  
 |||||  
 DB 426 CEKSFCK 434

RESULT 4  
 ID Q56914 PRELIMINARY; PRT; 318 AA.  
 AC Q56914;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE TRSB.  
 GN TRSB.  
 OS Versinia enterocolitica.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Versinia.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=6471/76 (SEROTYPE O:3);  
 RX MEDLINE; 96100456  
 RA SKURNIK M., VENHO R., TOIVANEN P., AL HENDY A.;  
 RT "A novel locus of *Versinia enterocolitica* serotype O:3 involved in  
 RL lipopolysaccharide outer core biosynthesis.";  
 RL Mol. Microbiol. 17:575-594(1995).  
 DR EMBL; Z47767; CAA87699.1; -;  
 DR PFAM; PF00535; Glycos\_transf\_2; 1.  
 SQ SEQUENCE 318 AA; 37112 MW; 7ADB7E51 CRC32;

Query Match 63.8%; Score 37; DB 2; Length 318;  
 Best Local Similarity 62.5%; Pred. No. 15;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCK 8  
 |||||  
 DB 193 CKSFECK 200

RESULT 5  
 ID Q74853 PRELIMINARY; PRT; 1077 AA.  
 AC Q74853;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE PUTATIVE CYSTINE-RICH TRANSCRIPTIONAL REGULATOR.  
 GN SPC18.03.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972H-;  
 RA HILBERT H., DUESTERHOEF A., WOOD V., RAJANDREAM M.A., BARRELL B.G.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031907; CAA21417.1; -;  
 DR PFAM; PF01424; R3H; 1.  
 DR PFAM; PF01422; zf-NF-X1; 8.  
 SQ SEQUENCE 1077 AA; 121087 MW; 99CE9378 CRC32;

Query Match 63.8%; Score 37; DB 3; Length 1077;  
 Best Local Similarity 66.7%; Pred. No. 37;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKLFCK 9  
 |||||  
 DB 702 CNKLSCK 710

RESULT 6  
 ID Q26055 PRELIMINARY; PRT; 1051 AA.  
 AC Q26055;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE 2-ALPHA COLLAGEN (COLL2-ALPHA) (FRAGMENT).  
 OS Paracentrotus lividus (Common sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea;  
 OC Euechinoidea; Echinacea; Echinoida; Echinidae; Paracentrotus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90216744.  
 RA D'ALESSIO M., RAMIREZ F., SUZUKI H.R., SOLURSH M., GAMBINO R.;  
 RT "Cloning of a fibrillar collagen gene expressed in the mesenchymal  
 RL cells of the developing sea urchin embryo.";  
 RL J. Biol. Chem. 265:7050-7054(1990).  
 DR EMBL; J05422; AAA29440.1; -;

DR PFAM; PF01410; COLFI; 1.  
 DR PFAM; PF01391; Collagen; 12.  
 FT NON\_TER 1 1  
 FT CHAIN 1 786 POTENTIAL.  
 SQ SEQUENCE 1051 AA; 101961 MW; 11D17FBD CRC32;

Query Match 63.8%; Score 37; DB 5; Length 1051;  
 Best Local Similarity 85.7%; Pred. No. 36;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKLFLFKC 7  
 |||||

Db 854 CKDLFLFKC 860

RESULT 7  
 O18387 PRELIMINARY; PRT; 774 AA.

AC O18387;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE D19B.  
 GN D19B.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
 OC Drosophilidae; Drosophila.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA JULLIEN D., KAS E.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ000261; CAA03977.1;  
 DR FLYBASE; FBgn0022699; D19B.  
 DR PFAM; PF00036; zf-C2H2; 12.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 10.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 774 AA; 88059 MW; D52BE4AF CRC32;

Query Match 63.8%; Score 37; DB 5; Length 774;  
 Best Local Similarity 55.6%; Pred. No. 29;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKLFLFKCT 9  
 |||||

Db 352 CEKVRCP 360

RESULT 8  
 Q26639 PRELIMINARY; PRT; 3198 AA.

AC Q26639;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE ALPHA-2 COLLAGEN.  
 GN COLP2ALPHA.  
 OS Strongylocentrotus purpuratus (Purple sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea;  
 OC Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;  
 OC Strongylocentrotus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92348411.  
 RA EXPOSITO J.Y., D'ALESSIO M., SOLURSH M., RAMIREZ F.;  
 RT "Sea urchin collagen evolutionarily homologous to vertebrate  
 pro-alpha 2(I) collagen."  
 RL J. Biol. Chem. 267:15559-15562(1992).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92381062.

RA EXPOSITO J.Y., D'ALESSIO M., RAMIREZ F.;  
 RT "Novel amino-terminal propeptide configuration in a fibrillar  
 procollagen undergoing alternative splicing."  
 RL J. Biol. Chem. 267:17404-17408(1992).  
 DR EMBL; M92041; AAA30040.1; -;  
 DR PFAM; PF01410; COLFI; 1.  
 DR PFAM; PF01391; Collagen; 17.  
 DR PFAM; PF00093; vwc; 1.  
 SQ SEQUENCE 3198 AA; 331465 MW; 9EF47D37 CRC32;

Query Match 63.8%; Score 37; DB 5; Length 3198;  
 Best Local Similarity 85.7%; Pred. No. 84;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKLFLFKC 7  
 |||||

Db 3002 CKDLFLFKC 3008

RESULT 9

Q86519 PRELIMINARY; PRT; 999 AA.  
 AC Q86519;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
 DE HYPOTHETICAL 110.5 KD PROTEIN.  
 OS Rice yellow mottle virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Sobemovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NIGERIA;  
 RA PINTO Y.M., BAULCOMBE D.C.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U23142; AAA64287.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 999 AA; 110470 MW; 8085C861 CRC32;

Query Match 63.8%; Score 37; DB 12; Length 999;  
 Best Local Similarity 50.0%; Pred. No. 35;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKLFLFKCTK 10  
 ||: |||:

Db 869 CKEYPCCTR 878

RESULT 10

Q86525 PRELIMINARY; PRT; 999 AA.  
 AC Q86525;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE POLYPROTEIN.  
 OS Rice yellow mottle virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Sobemovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94157480.  
 RA YASSI M.N., RITZENTHALER C., BRUGIDOU C., FAUQUET C., BEACHY R.N.;  
 RT "Nucleotide sequence and genome characterization of rice yellow  
 mottle virus RNA."  
 RL J. Gen. Virol. 75:249-257(1994).  
 DR EMBL; L20893; AAA72030.1; -;  
 KW Polyprotein.  
 SQ SEQUENCE 999 AA; 110686 MW; E3D12FCB CRC32;

Query Match 63.8%; Score 37; DB 12; Length 999;  
 Best Local Similarity 50.0%; Pred. No. 35;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 10  
| | : | | |  
Db 869 CKKYPCKTR 878

## RESULT 11

Q94993 ID Q94993 PRELIMINARY; PRT; 200 AA.  
AC Q94993;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE RIBOSOMAL PROTEIN S6 GENE AND TWO POTENTIAL ALTERNATIVELY SPLICED  
DE PROTEINS, COMPLETE CDS.  
GN RPS6.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
OC Drosophilidae; Drosophila.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93204998.  
RA STEWART M.J., DENELL R.;  
RT "Mutations in the Drosophila gene encoding ribosomal protein S6 cause  
RT tissue overgrowth.";  
RL Mol. Cell. Biol. 13:2524-2535(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94018587.  
RA STEWART M.J., DENELL R.;  
RT "The Drosophila ribosomal protein S6 gene includes a 3' triplication  
RT that arose by unequal crossing-over.";  
RL Mol. Biol. Evol. 10:1041-1047(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA STEWART M.J.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L02074; AAB05983.1; -;  
DR EMBL; L02074; AAB05984.1; -;  
DR FLYBASE; FBgn0004922; Rps6.  
DR PFAM; PF01092; Ribosomal\_S6; 1.  
DR PROSITE; PS00578; RIBOSOMAL\_S6; 1.  
SQ SEQUENCE 200 AA; 23072 MW; DC870C3B CRC32;

Query Match 62.1%; Score 36; DB 5; Length 200;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 9  
| | : | | |  
Db 87 CNKVRCKTK 95

## RESULT 12

O43279 ID O43279 PRELIMINARY; PRT; 1087 AA.  
AC O43279;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DE DSIF P160.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA WADA T., TAKAGI T., YAMAGUCHI Y., FERDOUS A., IMAI T., HIROSE S.,  
RA SUGIMOTO S., YANO K., HARTZOG G.A., WINSTON F., BURATOWSKI S.,  
RA HANDA H.;  
RL Genes Dev. 0:0-0(1998).

DR EMBL; AB000516; BAA24075.1; -;  
SQ SEQUENCE 1087 AA; 121058 MW; C81FCEAC CRC32;

Query Match 60.3%; Score 35; DB 4; Length 1087;  
Best Local Similarity 50.0%; Pred. No. 85;  
Matches 8; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

QY 1 CKKL-----FKCKTK 10  
| | | | : | | | |  
Db 626 CKKLVGGMFVCKTR 641

## RESULT 13

Q15911 ID Q15911 PRELIMINARY; PRT; 3703 AA.  
AC Q15911;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE ZINC FINGER HOMEODOMAIN PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92049333.  
RA MORINAGA T., YASUDA H., HASHIMOTO T., HIGASHIO K., TAMAOKI T.;  
RT "A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains  
RT four homeodomains and seventeen zinc fingers.";  
RL Mol. Cell. Biol. 11:6041-6049(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96070776.  
RA MIURA Y., TAM T., IDO A., MORINAGA T., MIKI T., HASHIMOTO T.,  
RA TAMAOKI T.;  
RT "Cloning and characterization of an ATBF1 isoform that expresses in a  
RT neuronal differentiation-dependent manner.";  
RL J. Biol. Chem. 270:26840-26848(1995).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL; L32832; AAC14462.1; -;  
DR PFAM; PF00046; homeobox; 4.  
DR PROSITE; PS00096; zf-C2H2; 20.  
DR PROSITE; PS00027; HOMEBOX\_1; 2.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 14.  
KW Homeobox; DNA-binding; Nuclear protein; Zinc-finger; Metal-binding.  
FT VARIANT 3528 3529 SY -> AA.  
SQ SEQUENCE 3703 AA; 404468 MW; 797F7644 CRC32;

Query Match 60.3%; Score 35; DB 4; Length 3703;  
Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKTK 9  
| | : | | |  
Db 2716 CRLFKAKT 2724

## RESULT 14

O60765 ID O60765 PRELIMINARY; PRT; 605 AA.  
AC O60765;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE HKL1.  
GN HKL1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.



RN SEQUENCE FROM N.A.  
 RP TISSUE=FETAL BRAIN;  
 RX MEDLINE; 98127052.  
 RA OMORI Y., KYUSHIKI H., TAKEDA S., SUZUKI M., KAWAI A., FUJIWARA T.,  
 RA TAKAHASHI E., NAKAMURA Y.,  
 RT "Cloning, expression and mapping of a novel human zinc-finger gene  
 RT TCF17 homologous to rodent Kif1.";  
 RL Cytogenet. Cell Genet. 78:285-288(1997).  
 DR EMBL; D89928; BAA25182.1; -.  
 DR PFAM; PF01352; KRAB; 1.  
 DR PFAM; PF00096; zf-C2H2; 13.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 13.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 605 AA; 69252 MW; BB70EF8F CRC32;

Query Match 60.3%; Score 35; DB 4; Length 605;  
 Best Local Similarity 85.7%; Pred. No. 54;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLKCK 8

DB 239 EKLKCK 245

RESULT 15

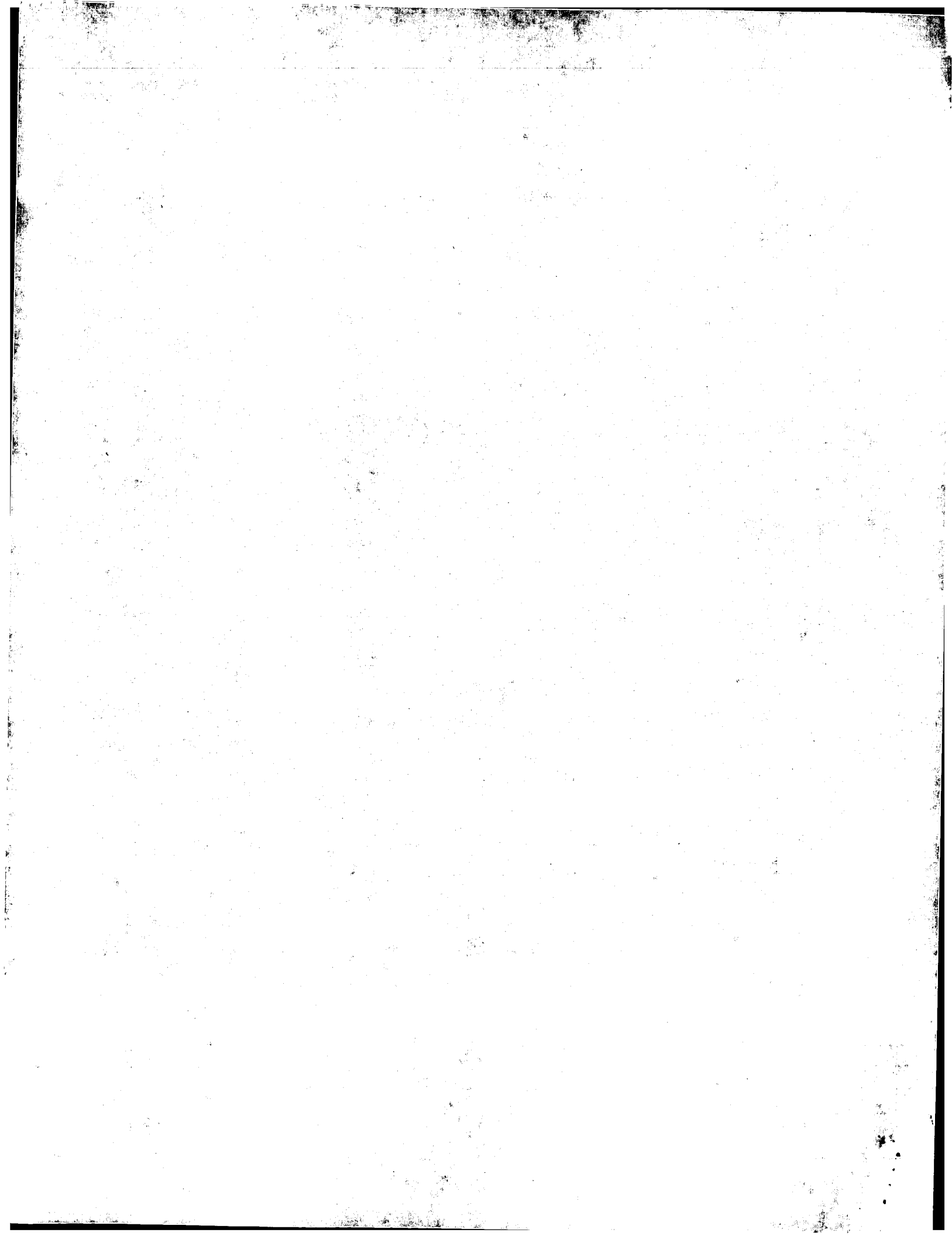
Q13719 PRELIMINARY; PRT; 2783 AA.  
 AC Q13719;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE ALPHA-FETOPROTEIN ENHANCER BINDING PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEPATOMA;  
 RX MEDLINE; 92049333.  
 RA MORINAGA T., YASUDA H., HIGASHIO K., TAMAOKI T.;  
 RT "A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains  
 RT four homeodomains and seventeen zinc fingers.";  
 RL Mol. Cell. Biol. 11:6041-6049(1991).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; D10250; BAA01095.1; -.  
 DR PFAM; PF00046; homeobox; 4.  
 DR PFAM; PF00096; zf-C2H2; 16.  
 DR PROSITE; PS00027; HOMEBOX\_1; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 11.  
 KW Homeobox; DNA-binding; Nuclear protein; Zinc-finger; Metal-binding.  
 FT VARIANT 2460 2460 A -> V.  
 SQ SEQUENCE 2783 AA; 305737 MW; DC7AAE23 CRC32;

Query Match 60.3%; Score 35; DB 4; Length 2783;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLKCKT 9

DB 1802 CRLKFKAKT 1810

Search completed: September 7, 1999, 22:47:45  
 Job time: 7965 sec





PR 11-FEB-1991: US-658744.  
PA (PORR/) PORRO M.

DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 7; Page 32; 39pp; English.

CC This peptide is a specific example of a generic peptide of formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where n = 1-100 and each R is H, an amino acid residue or a fatty acid residue. The peptide is useful for treating or preventing septic shock, mixing with polymyxin B to reduce its toxicity; removing endotoxins from blood, sera or other fluids (in vivo or in vitro); controlling release of cytokines induced by endotoxins; as diagnostic reagents to detect and quantify toxins in blood or sera; preparing non-toxic antigenic complexes of lipid A or lipopolysaccharide (LPS); and for treating pertussis, bacterial meningitis and HIV-related infections. The usual dose is 10-100 ug/kg/day, given parenterally. It binds to the same sites as polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It has no antibiotic activity; does not lyse erythrocytes; has no toxicity in mice when injected at 50mg/kg and is relatively unstable against proteases.

CC Sequence 9 AA;  
SQ

Query Match 100.0%; Score 53; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKCT 9  
| | | | | | | | | |  
DB 1 CKKLFCKCT 9

RESULT 3

R39290  
ID R39290 standard; peptide; 10 AA.

AC R39290;  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
KW quantification.

OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..7

PN WO9314115-A.  
PD 22-JUL-1993.

PF 14-MAY-1992; E01060.  
PR 16-JAN-1992; US-819893.

PA (PORR/) PORRO M.  
PI Porro M;

DR WPI; 93-243143/30.

PT New peptide(s) which neutralise lipid A of bacterial endotoxin - forming non-toxic, antigenic complex, used to treat or prevent septic shock, in vaccines to detoxify blood, etc.

PS Claim 6; Page 32; 45pp; English.

CC The sequence is that of a peptide which binds to the lipid A component of bacterial endotoxin at the same site as polymyxin B (PMB) and with about the same affinity to produce a non-toxic antigenic complex. Unlike PMB it is not toxic, is susceptible to proteolytic degradation in serum, has no antibiotic activity and no haemolytic action. It is especially used to treat or prevent septic shock; to reduce toxicity of PMB; to remove endotoxins from blood, sera, vaccines, drug solns. etc.; to control release of cytokines induced by endotoxins; for in vivo or in vitro detoxification of bacterial endotoxins; and to detect or quantify endotoxins in blood products.

CC Sequence 10 AA;  
SQ

Query Match 100.0%; Score 53; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKCT 9  
| | | | | | | | | |  
DB 1 CKKLFCKCT 9

RESULT 4

R39291  
ID R39291 standard; peptide; 9 AA.

AC R39291;  
DE 22-DEC-1993 (first entry)  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
KW quantification.

OS Synthetic.

FH Key Location/Qualifiers  
FT disulfide\_bond 1..7

PN WO9314115-A.  
PD 22-JUL-1993.

PF 14-MAY-1992; E01060.  
PR 16-JAN-1992; US-819893.

PA (PORR/) PORRO M.  
PI Porro M;

DR WPI; 93-243143/30.

PT New peptide(s) which neutralise lipid A of bacterial endotoxin - forming non-toxic, antigenic complex, used to treat or prevent septic shock, in vaccines to detoxify blood, etc.

PS Claim 7; Page 32; 45pp; English.

CC The sequence is that of a peptide which binds to the lipid A component of bacterial endotoxin at the same site as polymyxin B (PMB) and with about the same affinity to produce a non-toxic antigenic complex. Unlike PMB it is not toxic, is susceptible to proteolytic degradation in serum, has no antibiotic activity and no haemolytic action. It is especially used to treat or prevent septic shock; to reduce toxicity of PMB; to remove endotoxins from blood, sera, vaccines, drug solns. etc.; to control release of cytokines induced by endotoxins; for in vivo or in vitro detoxification of bacterial endotoxins; and to detect or quantify endotoxins in blood products.

CC Sequence 9 AA;  
SQ

Query Match 100.0%; Score 53; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKCT 9  
| | | | | | | | | |  
DB 1 CKKLFCKCT 9

RESULT 5

W21621  
ID W21621 standard; peptide; 10 AA.

AC W21621;  
DE 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #33.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
KW Synthetic.

FH Key Location/Qualifiers  
FT disulfide\_bond 1..7

PN WO9638163-A1.  
PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

```

PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M;
DR WPI: 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 37; Page 28; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to
CC potentiate the activity of an antibiotic when they are co-administered
CC with the antibiotic. Compositions containing these peptides are used
CC to treat or prevent microbial infections. These peptides bind to
CC lipopolysaccharide on the bacteria so may increase permeability of
CC the outer bacterial membrane to the antibiotic, allowing a reduction
CC in the dose of antibiotic required by 10-90% of the normal dose for
CC in vivo or in vitro application. Any toxic side effects are
CC correspondingly reduced.
SQ Sequence 10 AA;

Query Match 100.0%; Score 53; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKCT 9
DB 1 CKKLFCKCT 9

RESULT 6
W21622
ID W21622 standard; peptide; 9 AA.
AC W21622;
DT 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #34.
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;
KW permeability; outer bacterial membrane.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide_bond 1..7
FT WO9638163-A1.
PN 05-DEC-1996.
PF 29-MAY-1996; E02313.
PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M;
DR WPI: 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 38; Page 28; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to
CC potentiate the activity of an antibiotic when they are co-administered
CC with the antibiotic. Compositions containing these peptides are used
CC to treat or prevent microbial infections. These peptides bind to
CC lipopolysaccharide on the bacteria so may increase permeability of
CC the outer bacterial membrane to the antibiotic, allowing a reduction
CC in the dose of antibiotic required by 10-90% of the normal dose for
CC in vivo or in vitro application. Any toxic side effects are
CC correspondingly reduced.
SQ Sequence 9 AA;

Query Match 100.0%; Score 53; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKCT 9
DB 1 CKKLFCKCT 9

RESULT 7
R23962
ID R23962 standard; Protein; 2783 AA.

```

```

AC R23962;
DT 06-JAN-1993 (first entry)
DE AFP-1.
KW Enhancer; alpha fetoprotein; homeodomain; zinc-finger.
OS Homo sapiens.
FH Key Location/Qualifiers
FT domain 1231..1291
FT domain /note="homeodomain"
FT domain 1328..1388
FT domain /label="homeodomain"
FT domain 1727..1786
FT domain /label="homeodomain"
FT domain 2032..2091
FT domain /label="homeodomain"
FT domain 67..99
FT domain /label="zinc-finger"
FT domain 123..155
FT domain /label="zinc-finger"
FT domain 171..203
FT domain /label="zinc-finger"
FT domain 306..366
FT domain /label="zinc-finger"
FT domain 443..476
FT domain /label="zinc-finger"
FT domain 484..543
FT domain /label="zinc-finger"
FT domain 628..660
FT domain /label="zinc-finger"
FT domain 679..711
FT domain /label="zinc-finger"
FT domain 1066..1097
FT domain /label="zinc-finger"
FT domain 1411..1442
FT domain /label="zinc-finger"
FT domain 1613..1643
FT domain /label="zinc-finger"
FT domain 1794..1825
FT domain /label="zinc-finger"
FT domain 2107..2139
FT domain /label="zinc-finger"
FT domain 2540..2571
FT domain /label="zinc-finger"
FT domain 2606..2638
FT domain /label="zinc-finger"
PN EP-487229-A.
PD 27-MAY-1992.
PF 07-NOV-1991; 310334.
PR (SNOW) SNOW BRAND MILK PROD CO LTD.
PA Higashio K, Morinaga T, Tamaoki T, Yasuda N;
DR WPI: 92-176828/22.
DR N-PSDB; Q24828.
PT DNA encoding protein binding to alpha-fetoprotein gene enhancer -
PT useful for prodn. of biological active protein
PS Claim 1; Page 7; 24pp; English.
CC This protein specifically binds to the enhancer of the alpha-
CC fetoprotein gene. The DNA encoding this protein may be useful for
CC the prodn. of biologically active proteins by inserting the DNA into
CC an expression vector and co-transfecting animal cells with another
CC expression vector in which a gene for the biologically active protein
CC is placed under the control of alpha-fetoprotein gene enhancer and
CC promoter.
SQ Sequence 2783 AA;

Query Match 66.0%; Score 35; DB 1; Length 2783;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKCT 9
DB 1802 CRALFRKKT 1810

```

RESULT 8  
 R23963  
 ID R23963 standard; Protein; 2783 AA.  
 AC R23963;  
 DT 06-JAN-1993 (first entry)  
 DE AFP-1 (Ala 2460 Val).  
 KW Enhancer; alpha fetoprotein; homeodomain; zinc-finger.  
 OS Homo sapiens.  
 FH Key  
 FT Location/Qualifiers  
 FT 1231..1291  
 FT /note= "homeodomain"  
 FT 1328..1388  
 FT /label= homeodomain  
 FT 1727..1786  
 FT /label= homeodomain  
 FT 2032..2091  
 FT /label= homeodomain  
 FT 67..99  
 FT /label= zinc-finger  
 FT 123..155  
 FT /label= zinc-finger  
 FT 171..203  
 FT /label= zinc-finger  
 FT 306..366  
 FT /label= zinc-finger  
 FT 443..476  
 FT /label= zinc-finger  
 FT 484..543  
 FT /label= zinc-finger  
 FT 628..660  
 FT /label= zinc-finger  
 FT 679..711  
 FT /label= zinc-finger  
 FT 1066..1097  
 FT /label= zinc-finger  
 FT 1411..1442  
 FT /label= zinc-finger  
 FT 1613..1643  
 FT /label= zinc-finger  
 FT 1794..1825  
 FT /label= zinc-finger  
 FT 2107..2139  
 FT /label= zinc-finger  
 FT 2540..2571  
 FT /label= zinc-finger  
 FT 2606..2638  
 FT /label= zinc-finger  
 PN EP-487229-A.  
 PD 27-MAY-1992.  
 PF 07-NOV-1991; 310334.  
 PR (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Higashio K, Morinaga T, Tamaoki T, Yasuda N;  
 DR WPI: 92-176828/22.  
 DR N-PSDB: Q24829.  
 PT DNA encoding protein binding to alpha-fetoprotein gene enhancer -  
 PT useful for prodn. of biological active protein  
 PS Claim 3-4; Page 7; 24pp; English.  
 CC This protein sequence is as the sequence given in R23962 except  
 CC that Ala at position 2460 is replaced by Val.  
 CC This protein specifically binds to the enhancer of the alpha-  
 CC fetoprotein gene. The DNA encoding this protein may be useful for  
 CC the prodn. of biologically active proteins by inserting the DNA into  
 CC an expression vector and co-transfecting animal cells with another  
 CC expression vector in which a gene for the biologically active protein  
 CC is placed under the control of alpha-fetoprotein gene enhancer and  
 CC promoter.  
 SQ Sequence 2783 AA;

Query Match 66.0%; Score 35; DB 1; Length 2783;  
 Best Local Similarity 66.7%; Pred. No. 4e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 CKLFLCKCT 9  
 I: ||| ||  
 Db 1802 CRALFKAKT 1810  
 RESULT 9  
 R06646  
 ID R06646 standard; protein; 551 AA.  
 AC R06646;  
 DT 07-JAN-1991 (first entry)  
 DE Human interleukin-2 receptor beta-chain.  
 KW IL-2R beta; PIL-2R beta 9.  
 OS Homo sapiens.  
 PN EP-386304-A.  
 PD 12-SEP-1990.  
 PF 29-MAY-1989; 109656.  
 PR 07-MAR-1989; EP-104023.  
 PR 29-MAY-1989; EP-109656.  
 PA (OSAU ) OSAKA UNIVERSITY.  
 PT Taniiguchi T;  
 DR WPI: 90-276456/37.  
 DR N-PSDB: Q05872.  
 PT Recombinant interleukin-2 receptor beta chain - used for studying  
 PT IL-2 system and producing antibodies for diagnosis and therapy.  
 PS Claim 4; Fig 1B; 37pp; English.  
 CC IL-2 receptor beta chains are useful in studying the biochemistry of  
 CC interleukin, and in diagnosis and therapy by immune suppression and  
 CC activation.  
 SQ Sequence 551 AA;

Query Match 62.3%; Score 33; DB 1; Length 551;  
 Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KKLFLCKCT 9  
 ||| ||| |  
 Db 275 KKLKRCNT 282  
 RESULT 10  
 R07506  
 ID R07506 standard; protein; 551 AA.  
 AC R07506;  
 DT 07-FEB-1991 (first entry)  
 DE IL-2R beta chain.  
 KW Interleukin; receptor.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..26  
 FT /label=sol.IL-2R beta  
 FT 27..551  
 FT /label=transmembrane region  
 FT 241..265  
 FT /label=N-glycos\_site  
 FT 29..31  
 FT /label=N-glycos\_site  
 FT 43..45  
 FT /label=N-glycos\_site  
 FT 71..73  
 FT /label=N-glycos\_site  
 FT 149..151  
 FT /label=N-glycos\_site  
 FT 352..354  
 FT /label=N-glycos\_site  
 PN AU9050726-A.  
 PD 13-SEP-1990.  
 PF 06-MAR-1990; 050726.  
 PR 07-MAR-1989; EP-104023.  
 PR 29-MAY-1989; EP-109656.  
 PR 20-JUL-1989; EP-113310.

PA (BOEH ) BOEHRINGER INGELHEI.  
 PI Taniguchi T, Hatakeyama M, Minamoto S, Kono T, Doi T;  
 PI Miyasaka M, Tsudo, Karasuyama H;  
 DR WPI: 90-327673/44.  
 DR N-PSDB: Q06331.  
 PT Recombinant interleukin-2 receptor beta chain - useful for  
 PT diagnosis and therapy by immune suppression or activation.  
 PS Claim 3; Fig 1B; 63pp; English.  
 CC The sequence was deduced from cDNA obt'd. from clones isolated from  
 CC a cDNA library prepd. from RNA extracted from the NK-like human  
 CC lymphoid cell line, YT. The clones were selected using a cocktail  
 CC of anti-IL-2R beta MAbs, Mik-beta1 and Mik-beta2. The sequence  
 CC encodes a soluble portion of the IL-2R beta chain.  
 CC See also R07507.  
 SQ Sequence 551 AA;

Query Match 62.3%; Score 33; DB 1; Length 551;  
 Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KKLCKCT 9  
 III III  
 DB 275 KLLKCN 282

RESULT 11  
 W19786 ID W19786 standard; Protein; 1088 AA.  
 AC W19786;  
 DT 08-NOV-1997 (first entry)  
 DE Human multiple regulatory protein SPT5.  
 KW SPT5; multiple function regulatory protein; MFRP; cell cycle;  
 KW gastrulation; carcinogenesis; tumorigenicity; breast cancer;  
 KW tumour suppressor; neoplasm; carcinogenesis; organogenesis;  
 KW gonad development.  
 OS Homo sapiens.

FH Key Location/Qualifiers  
 FT peptide 827..830  
 FT /note= "MAPK consensus site"  
 FT peptide 902..905  
 FT /note= "MAPK consensus site"  
 FT peptide 933..937  
 FT /note= "MAPK consensus site"  
 FT peptide 944..948  
 FT /note= "MAPK consensus site"  
 FT peptide 953..956  
 FT /note= "MAPK consensus site"  
 FT peptide 966..969  
 FT /note= "MAPK consensus site"  
 FT peptide 1033..1036  
 FT /note= "MAPK consensus site"

WO9703087-A1.  
 PN 30-JAN-1997.  
 PD 12-JUL-1996; U11702.  
 PF 13-JUL-1995; US-001113.  
 PR (UNMI ) UNIV MICHIGAN.  
 PI Chiang P, Kurnit DM, Song W;  
 DR WPI: 97-118981/11.  
 DR N-PSDB: T72814.  
 PT New family of multiple function regulatory proteins - regulate cell  
 PT cycle progression, gastrulation, carcinogenesis, tumorigenicity,  
 PT gonad development, etc.  
 PS Claim 32; Fig 14b; 98pp; English.  
 CC This polypeptide is a novel multiple function regulatory protein  
 CC (MFRP) that shows homology to yeast SPT5. It is encoded by a cDNA  
 CC clone (T72814) obt'd. from a human lung cDNA library. Mammalian  
 CC SPT5, SPT4 and SPT6 (see also W19782-85) homologues of yeast SPT  
 CC transcription factors regulate cellular functions such as cell  
 CC cycle progression, control of gastrulation, carcinogenesis (such as  
 CC breast cancer), tumorigenicity, inhibition or reversal of a  
 CC neoplastic phenotype, progression through the cell cycle, gonad  
 CC development, carcinogenesis and organogenesis. They can be

CC expressed in transformed host cells and used to generate antibodies  
 CC used in a claimed method for detecting wild-type mammalian MFRP, or  
 CC expressed using gene therapy methods to restore wild-type MFRP  
 CC function, such as tumour suppression.  
 SQ Sequence 1088 AA;

Query Match 62.3%; Score 33; DB 1; Length 1088;  
 Best Local Similarity 53.3%; Pred. No. 3.8e+02;  
 Matches 8; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 CKKL-----FKCKT 9  
 IIII I III  
 DB 627 CKKLVENGMFVCKT 641

RESULT 12  
 W88480 ID W88480 standard; Protein; 971 AA.  
 AC W88480;  
 DT 10-MAY-1999 (first entry)  
 DE Candida albicans histidine kinase 1 CAHK1 partial sequence.  
 KW CAHK1 gene; histidine kinase-1; two-component gene;  
 KW signal transduction; virulence; candidosis; diagnosis; therapy;  
 KW vaccine; infection.  
 OS Candida albicans.  
 FH Key Location/Qualifiers  
 FT Domain 482..721  
 FT /note= "Sensor domain, this polypeptide is  
 FT specifically claimed in Claim 13(g)"  
 FT Domain 834..971  
 FT /note= "response regulator domain, this polypeptide  
 FT is specifically claimed in Claim 13(h)"  
 FT Protein 2..971  
 FT /note= "this polypeptide region is specifically  
 FT claimed in Claim 13(c)"  
 FT Protein 482..971  
 FT /note= "this polypeptide region is specifically  
 FT claimed in Claim 13(i)"

WO9902700-A1.  
 PN 21-JAN-1999.  
 PD 09-JUL-1998; U14254.  
 PF 11-FEB-1998; US-074308.  
 PR 10-JUL-1997; US-052273.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (GEOU ) UNIV GEORGETOWN.  
 PI Calderone RA, Calera JA, Choi GH;  
 DR WPI: 99-120901/10.  
 DR N-PSDB: X06981.  
 PT New isolated Candida albicans Histidine Kinase-1 gene - used to  
 PT develop products for the diagnosis, prevention and treatment of C.  
 PT albicans infections, particularly oral candidosis  
 PS Claim 13(a); Page 86-89; 106pp; English.  
 CC This polypeptide comprises a portion of Candida albicans  
 CC histidine kinase-1 (CAHK1), and is encoded by a partial length  
 CC DNA clone (see X06981) deposited as ATCC 209504. The full-length  
 CC 2471-amino acid CAHK-1 protein (282 kDa) is provided in W88481.  
 CC CAHK1 is active in phosphorylating host cell proteins to render  
 CC the host cell susceptible to invasion by C. albicans, i.e. is  
 CC involved in virulence. It exhibits homology with the sensor and  
 CC regulator components of prokaryotic and eukaryotic two-component  
 CC histidine kinases. Also provided are vectors, host cells,  
 CC antibodies and recombinant methods for producing them. The  
 CC invention also provides agonists and antagonists of CAHK1 activity,  
 CC as well as diagnostic methods for detecting CAHK1 nucleic acids,  
 CC polypeptides and antibodies in a biological sample. CAHK1  
 CC antagonists can be used to treat C. albicans infections, e.g. oral  
 CC candidosis, including oral candidosis associated with HIV  
 CC infection. The products can also be used in e.g. immunoassays, in  
 CC epitope mapping, epitope tagging, to generate antibodies to a  
 CC particular portion of the protein, and as vaccines.  
 SQ Sequence 971 AA;

Query Match 62.3%; Score 33; DB 1; Length 971;  
 Best Local Similarity 55.6%; Pred. No. 3.4e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKT 9  
 DB 7 CSQHFCKS 15

RESULT 13  
 ID R92103 standard; Protein; 1706 AA.  
 AC R92103;  
 DT 19-MAY-1996 (first entry)  
 DE Rat RIZ.  
 KW Retinoblastoma protein-interacting zinc finger; RIZ;  
 KW cell proliferation; tumour; cancer; neuroblastoma; melanoma;  
 KW diagnosis; therapy.  
 OS Rattus sp.

FT Key Location/Qualifiers

FT domain 36..151

FT domain /label= PR\_domain

FT domain 304..309

FT region /label= cr2\_core\_motif

FT region 357..377

FT region /label= zinc\_finger(1)

FT region 387..407

FT region /label= zinc\_finger(2)

FT region 478..499

FT region /label= zinc\_finger(3)

FT region 667..695

FT region /label= Leucine-zipper

FT domain 867..874

FT region /label= Nuclear\_localisation\_signal

FT region 1125..1203

FT region /label= zinc\_finger(4-6)

FT region 1323..1343

FT region /label= zinc\_finger(7)

FT region 1445..1466

FT region /label= zinc\_finger(8)

PN W0906168-A2.

PD 29-FEB-1996.

PE 18-AUG-1995; U10574.

PR 18-AUG-1994; US-292683.

PR 06-MAR-1995; US-399411.

PA (LJOL-) LA JOLLA CANCER RES FOUND.

PI Huang S;

DR WPI; 96-151371/15.

DR N-PSDB; T18023.

PT Nucleic acid encoding mammalian retinoblastoma protein-interacting

PT zinc finger - used to identify cell proliferation modulating agents

PT for treatment of tumours, esp. neuroblastoma and melanoma, also used

PT for cancer diagnosis.

PS Claim 11; Fig 1A; 142pp; English.

CC The sequence (R92103) of full-length rat retinoblastoma (Rb)

CC protein-interacting zinc finger (RIZ) was deduced from a cDNA

CC sequence (T18023) cloned from a rat neonatal cardiac myocyte

CC library. RIZ is expressed in rat cells as a 250 kDa

CC phosphoprotein. It acts as a cell differentiation factor, and

CC regulates cell proliferation by binding to Rb, which is involved in

CC transcription. RIZ can also act to regulate

CC obtd. by recombinant DNA methods, and used to identify agents that

CC modulate cell proliferation and thereby treat tumour growth or

CC neurodegenerative disorders.

CC Sequence 1706 AA;

Query Match

Best Local Similarity 60.4%; Score 32; DB 1; Length 1706;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKT 9  
 DB 1156 CVOLEFKVT 1164

RESULT 14  
 W44777

ID W44777 standard; Protein; 1141 AA.

AC W44777;

DT 01-JUN-1998 (first entry)

DE Human Tbc-1 protein.

KW tre-2; BUB2; cdc16; mast cell; probe; hybridisation; subtraction method;

KW mouse; transcription factor; differentiation; proliferation; human;

KW acute myelogenous leukaemia.

OS Homo sapiens.

PN US5700927-A.

PD 23-DEC-1997.

PD 23-DEC-1994; 363300.

PR 23-DEC-1994; US-363300.

PA (CHIL-) CHILDRENS MEDICAL CENT.

PI Richardson P, Zon L;

DR WPI; 98-062437/06.

DR N-PSDB; V05886.

PT DNA encoding Tbc1 polypeptide - useful for treating leukaemia

PS Claim 1; Fig 1A-B; 22pp; English.

CC This is the amino acid sequence of a Tbc1 (tre-2, BUB2 and cdc16)

CC polypeptide, isolated from a bone marrow-derived mast cell DNA library.

CC The screen was carried out using a probe generated by a subtraction

CC method which compared mRNA expression in an undifferentiated mast cell

CC line P815 and cell line PGR6 (P815 cells transformed to express the

CC murine GATA-1 transcription factor - a factor which controls the

CC expression of genes involved in mast cell differentiation). Tbc1 encodes

CC a protein involved in the coupling of cell proliferation to cell

CC differentiation, which can be used to treat leukaemia (especially acute

CC myelogenous leukaemia) by causing leukaemic cells to differentiate.

SQ Sequence 1141 AA;

Query Match 60.4%; Score 32; DB 1; Length 1141;

Best Local Similarity 50.0%; Pred. No. 5.8e+02;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKT 8  
 DB 54 CSSIFECK 61

RESULT 15  
 R71789

ID R71789 standard; peptide; 10 AA.

AC R71789;

DT 01-OCT-1995 (first entry)

DE Peptide neutralising toxicity of Lipid A.

KW Endotoxin; Lipid A; septic shock; Lipopolysaccharide.

OS Synthetic.

FT Key Location/Qualifiers

FT disulfide bond 2..9

PN W09503327-A.

PD 02-FEB-1995.

PF 21-JUL-1994; E02413.

PR 26-JUL-1993; US-097830.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M;

DR WPI; 95-075190/10.

PT New peptide(s) for neutralising LPS endotoxin - comprising

PT repeating units of a basic amino acid or basic and hydrophobic

PT amino acids

PS Claim 20; Page 21; 26pp; English.

CC New peptides are claimed which are linear or cyclic peptides which

CC include units of formula: (A)n, where A is the cationic amino acid Lys

CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic

CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and

CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or



CC Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides.  
SQ Sequence 10 AA;

Query Match 59.4%; Score 31.5; DB 1; Length 10;  
Best Local Similarity 77.8%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CK-KLFRCK 8  
|||  
Db 2 CKFKKFRCK 10

Search completed: September 7, 1999, 22:50:06  
Job time: 7887 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:49 ; Search time 105.14 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-34

Perfect score: 53

Sequence: 1 CKKLFCKT 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	81.1	414	2 JQ1016	glycine hydroxymet
2	39	73.6	32	2 S29217	neurotoxin Tx2 - s
3	39	73.6	154	2 D39384	finger protein Hrf
4	39	73.6	803	2 S26823	finger protein ZNF
5	38	71.7	241	2 E70470	conserved hypotet
6	37	69.8	527	2 S53835	NADH dehydrogenase
7	37	69.8	402	2 A69002	glutamyl-tRNA redu
8	37	69.8	1051	2 A35763	collagen alpha 2 c
9	37	69.8	3198	2 A43426	collagen alpha 2 f
10	37	69.8	318	2 S51261	probable beta-glyc
11	35	66.0	2783	1 A41948	alpha-fetoprotein
12	35	66.0	576	2 A48157	renal transcriptio
13	34	64.2	420	2 C71634	serine hydroxymeth
14	34	64.2	704	2 F64373	hypothetical prote
15	34	64.2	328	2 G65110	hypothetical 35.2
16	34	64.2	118	2 C64441	conserved hypotet
17	34	64.2	116	2 S50449	hypothetical prote
18	34	64.2	770	2 S69075	probable membrane
19	34	64.2	204	2 S07741	hypothetical prote
20	34	64.2	86	2 I61885	zinc finger protei
21	34	64.2	189	2 F39384	finger protein Hrf
22	34	64.2	1191	2 S35305	zinc finger protei
23	33	62.3	522	1 KJHUG6	N-acetylglactosam
24	33	62.3	568	1 RRN239	polymerase-associa
25	33	62.3	568	1 RRN235	polymerase-associa
26	33	62.3	568	1 RRN273	polymerase-associa
27	33	62.3	568	1 RRN283	polymerase-associa
28	33	62.3	397	2 S42549	polygalacturonase
29	33	62.3	399	2 T00631	hypothetical prote
30	32	60.4	691	2 S78135	NADH dehydrogenase
31	32	60.4	889	2 S47162	DNA-directed RNA p
32	32	60.4	377	2 S77142	tryptophan--trNA 1
33	32	60.4	630	2 JQ1894	hypothetical prote
34	32	60.4	75	2 D64475	hypothetical prote
35	32	60.4	773	2 T01539	hypothetical prote
36	32	60.4	168	2 S06570	finger protein (cl
37	32	60.4	383	2 C32891	finger protein 9,
38	32	60.4	163	2 G39384	finger protein Hrf
39	32	60.4	732	2 S47073	finger protein H2F

40 32 60.4 39 2 B40984 finger protein zfe  
41 32 60.4 554 2 JU0188 membrane protein 4  
42 32 60.4 1706 2 I84499 zinc finger protei  
43 31 58.5 765 1 TVHUBF protein kinase B-r  
44 31 58.5 806 1 JN0612 protein kinase B-r  
45 31 58.5 807 1 I51153 protein kinase B-r

ALIGNMENTS

RESULT 1  
JQ1016  
glycine hydroxymethyltransferase (EC 2.1.2.1) - Campylobacter jejuni  
N: Alternate names: serine aldolase; serine hydroxymethylase; serine methylase; threon  
C: Species: Campylobacter jejuni  
C: Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 08-Sep-1997  
C: Accession: JQ1016; S16468  
R: Chan, V.L.; Bingham, H.L.  
Gene 101, 51-58, 1991  
A: Title: Complete sequence of the Campylobacter jejuni glyA gene encoding serine hydr  
A: Reference number: JQ1016; MUID: 91285434  
A: Accession: JQ1016  
A: Molecule type: DNA  
A: Residues: 1-414 <CHAI>  
A: Cross-references: EMBL:X53816; NID: g40533; PID: g40534  
R: Chan, V.L.; Bingham, H.L.  
Gene 101, 51-58, 1990  
A: Title: Complete sequence of the Campylobacter jejuni glyA gene encoding serine hydr  
A: Reference number: S16468  
A: Accession: S16468  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-414 <CHAI>  
A: Cross-references: EMBL:X53816; NID: g40533; PID: g40534  
C: Comment: In the presence of tetrahydrofolate, this enzyme catalyzes the reversible  
C: Genetics:  
A: Gene: glyA  
C: Superfamily: glycine hydroxymethyltransferase  
C: Keywords: transferase  
  
Query Match 81.1%; Score 43; DB 2; Length 414;  
Best Local Similarity 87.5%; Pred. No. 1.5;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 CKKLFCKK 8  
Db 77 CKKLFNCK 84  
  
RESULT 2  
S29217  
neurotoxin Tx2 - spider (Phonetrulia nigriventer)  
C: Species: Phonetrulia nigriventer  
C: Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Jul-1997  
C: Accession: S29217  
R: do Nascimento Cordeiro, M.; Ribeiro Diniz, C.; do Carmo Valentim, A.; von Eickstedt  
FEBS Lett. 310, 153-156, 1992  
A: Title: The purification and amino acid sequences of four Tx2 neurotoxins from the v  
A: Reference number: S29214  
A: Accession: S29217  
A: Status: preliminary  
A: Molecule type: protein  
A: Residues: 1-32 <COR>  
  
Query Match 73.6%; Score 39; DB 2; Length 32;  
Best Local Similarity 77.8%; Pred. No. 0.98;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 CKKLFCKCKT 9  
| | | | |

Db 15 CKKFKCKT 23

## RESULT 3

D39384  
finger protein HTF6 - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 10-Sep-1997  
C:Accession: D39384  
R:Bellet, E.J.; Poncellet, D.A.; Lecocq, P.J.; Revelant, O.; Martial, J.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 3608-3612, 1991  
A:Title: The evolutionarily conserved Krueppel-associated box domain defines a subfamily  
A:Reference number: A39384; MUID:91219421  
A:Accession: D39384  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-154 <BEL>  
A:Cross-references: GB:M61869; NID:9184449; PID:9184450  
A:Keywords: DNA binding; zinc; zinc finger

Query Match 73.6%; Score 39; DB 2; Length 154;  
Best Local Similarity 100.0%; Pred. No. 3.4; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0

QY 2 KKLKFKCK 8

|||||

Db 129 KKLKFKCK 135

## RESULT 4

S26823  
finger protein ZNF43 - human  
C:Species: Homo sapiens (man)  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Sep-1998  
C:Accession: S26823  
R:Lovering, R.; Trowsdale, J.  
Nucleic Acids Res. 19, 2921-2928, 1991  
A:Title: A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell line  
A:Reference number: S26823; MUID:91279444  
A:Accession: S26823  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-803 <LOV>  
A:Cross-references: EMBL:X59244; NID:938031; PID:938032  
C:Genetics:  
A:Gene: GDB:ZNF43; HTF6  
A:Cross-references: GDB:128653  
A:Map position: 19p13.1-19p12  
C:Keywords: DNA binding

Query Match 73.6%; Score 39; DB 2; Length 803;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLKFKCK 8

|||||

Db 165 KKLKFKCK 171

## RESULT 5

E70470  
conserved hypothetical protein aq\_1986 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C>Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 08-May-1998  
C:Accession: E70470  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: E70470

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-241 <AOF>  
A:Cross-references: GB:AE000766; NID:92984216; PID:92984233; GB:AE000657  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: aq\_1986

Query Match 71.7%; Score 38; DB 2; Length 241;  
Best Local Similarity 75.0%; Pred. No. 7.2;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCKK 8

|||||

Db 162 CKKLFCKK 169

## RESULT 6

S53835  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - Acanthamoeba castellanii mitoc  
C:Species: mitochondrion Acanthamoeba castellanii  
C>Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Feb-1999  
C:Accession: S53835  
R:Burger, G.; Plante, I.; Loneragan, K.M.; Gray, M.W.  
J. Mol. Biol. 245, 522-537, 1995  
A:Title: The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii: c  
A:Reference number: S53825; MUID:95147275  
A:Accession: S53835  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-527 <BUR>  
A:Cross-references: GB:U12386; NID:9562028; PID:9562039  
A:Experimental source: strain Neff; ATCC 30010  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C:Genetics:  
A:Genome: mitochondrion  
A:Genetic code: SGC6  
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 2  
C:Keywords: mitochondrion; NAD; oxidoreductase

Query Match 69.8%; Score 37; DB 2; Length 527;

Best Local Similarity 87.5%; Pred. No. 20;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKLKFKCKT 9

|||||

Db 13 KKLKFCGT 20

## RESULT 7

A69002  
glutamyl-tRNA reductase - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
C:Accession: A69002  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.  
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, K.  
; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu  
A:Reference number: A69000; MUID:98037514  
A:Accession: A69002  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-402 <MTH>  
A:Cross-references: GB:AE000874; GB:AE000666; NID:92622110; PID:92622114  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1012  
A:Start codon: GTG  
C:Superfamily: glutamyl-tRNA reductase

Query Match 69.8%; Score 37; DB 2; Length 402;  
 Best Local Similarity 75.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFKCK 8  
 I I I I I I I I  
 Db 392 CAKLFNCK 399

RESULT 8  
 A35763  
 collagen alpha 2 chain - sea urchin (Paracentrotus lividus) (fragment)  
 C:Species: Paracentrotus lividus (common urchin)  
 C:Date: 12-Oct-1990 #sequence\_revision 12-Oct-1990 #text\_change 08-Sep-1997  
 C:Accession: A35763  
 R: D'Alessio, M.; Ramirez, F.; Suzuki, H.R.; Solursh, M.; Gambino, R.  
 J. Biol. Chem. 265: 7050-7054, 1990  
 A:Title: Cloning of a fibrillar collagen gene expressed in the mesenchymal cells of the  
 A:Reference number: A35763; MUID:90216744  
 A:Accession: A35763  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1051 <DAA>  
 A:Cross-references: GB:J05422; NID:g159961; PID:g159962  
 C:Superfamily: unassigned collagens; fibrillar collagen carboxyl-terminal homology  
 F:830-1051/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 69.8%; Score 37; DB 2; Length 1051;  
 Best Local Similarity 85.7%; Pred. No. 34;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFKCK 7  
 I I I I I I I I  
 Db 854 CKDLFKC 860

RESULT 9  
 A43426  
 collagen alpha 2 fibrillar chain precursor - sea urchin (Strongylocentrotus purpuratus)  
 C:Species: Strongylocentrotus purpuratus (purple urchin)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
 C:Accession: A43426  
 R:Exposito, J.Y.; D'Alessio, M.; Ramirez, F.  
 J. Biol. Chem. 267, 17404-17408, 1992  
 A:Title: Novel amino-terminal propeptide configuration in a fibrillar procollagen undergo  
 A:Reference number: A43426; MUID:92381062  
 A:Accession: A43426  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-3198 <EXP>  
 A:Cross-references: GB:M92041; NID:g161448; PID:g161449  
 A:Note: sequence extracted from NCBI backbone (NCBIP:111965)  
 C:Superfamily: unassigned collagens; fibrillar collagen carboxyl-terminal homology; von  
 F:48-106/Domain: von Willebrand factor type C repeat homology <VWC>  
 F:2978-3198/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 69.8%; Score 37; DB 2; Length 3198;  
 Best Local Similarity 85.7%; Pred. No. 82;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFKCK 7  
 I I I I I I I I  
 Db 3002 CKDLFKC 3008

RESULT 10  
 S51261  
 Probable beta-glycosyltransferase trsb - Yersinia enterocolitica  
 C:Species: Yersinia enterocolitica

C:Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
 C:Accession: S51261; S70737  
 R:Skurnik, M.; Venho, R.; Toivanen, P.; Al Hendy, A.  
 submitted to the EMBL Data Library, January 1995  
 A:Description: A novel locus of Yersinia enterocolitica serotype O:3 involved in lipo  
 A:Reference number: S51258  
 A:Accession: S51261  
 A:Molecule type: DNA  
 A:Residues: 1-318 <SKU>  
 A:Cross-references: EMBL:247767; NID:g633689; PID:g633693  
 A:Experimental source: strain 6471/76 serotype O:3  
 R:Skurnik, M.; Venho, R.; Toivanen, P.; Al-Hendy, A.  
 Mol. Microbiol. 17, 575-594, 1995  
 A:Title: A novel locus of Yersinia enterocolitica serotype O:3 involved in lipopolysa  
 A:Reference number: S70734; MUID:96100456  
 A:Accession: S70737  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-110 <SKW>  
 A:Cross-references: EMBL:247767  
 A:Experimental source: strain 6471/76 serotype O:3  
 C:Genetics:  
 A:Gene: trsB  
 C:Keywords: lipopolysaccharide core biosynthesis

Query Match 69.8%; Score 37; DB 2; Length 318;  
 Best Local Similarity 62.5%; Pred. No. 13;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFKCK 8  
 I I I I I I I I  
 Db 193 CKSIFECK 200

RESULT 11  
 A41948  
 alpha-fetoprotein enhancer-binding protein - human  
 N:Alternate names: ATBF1 protein  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 14-Nov-1997  
 C:Accession: A41948  
 R:Morinaga, T.; Yasuda, H.; Hashimoto, T.; Higashio, K.; Tamaoki, T.  
 Mol. Cell. Biol. 11, 6041-6049, 1991  
 A:Title: A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four hom  
 A:Reference number: A41948; MUID:92049333  
 A:Accession: A41948  
 A:Molecule type: mRNA  
 A:Residues: 1-2783 <MOR>  
 A:Cross-references: GB:D10250; GB:D90395; NID:g219429; PID:g219430  
 A:Note: sequence extracted from NCBI backbone (NCBIN:66271, NCBIP:66276)  
 C:Genetics:  
 A:Gene: GDB:ATBF1  
 A:Cross-references: GDB:392090; OMIM:104155  
 A:Map position: 16q22.3-16q23.1  
 C:Superfamily: alpha-fetoprotein enhancer-binding protein; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription factor; zinc finger  
 F:73-94/Region: zinc finger CCHH motif  
 F:128-150/Region: zinc finger CCHH motif  
 F:176-198/Region: zinc finger CCHH motif  
 F:311-332/Region: zinc finger CCHH motif  
 F:340-361/Region: zinc finger CCHH motif  
 F:448-471/Region: zinc finger CCHH motif  
 F:489-509/Region: zinc finger CCHH motif  
 F:517-538/Region: zinc finger CCHH motif  
 F:633-655/Region: zinc finger CCHH motif  
 F:684-706/Region: zinc finger CCHH motif  
 F:719-773/Region: serine/threonine-rich  
 F:809-958/Region: glutamine-rich  
 F:1071-1092/Region: zinc finger CCHH motif  
 F:1117-1211/Region: proline-rich  
 F:1232-1288/Domain: homeobox homology <HOX1>  
 F:1329-1385/Domain: homeobox homology <HOX2>

F:1416-1437/Region: zinc finger CCH motif  
 F:1618-1638/Region: zinc finger CCH motif  
 F:1728-1784/Domain: homeobox homology <HOX3>  
 F:1799-1820/Region: zinc finger CCH motif  
 F:2033-2089/Domain: homeobox homology <HOX4>  
 F:2112-2134/Region: zinc finger CCH motif  
 F:2145-2366/Region: zinc finger CCH motif  
 F:2585-2607/Region: zinc finger CCH motif  
 F:2611-2633/Region: zinc finger CCH motif  
 F:2650-2737/Region: serine/threonine-rich

Query Match 66.0%; Score 35; DB 1; Length 2783;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKT 9

Db 1802 CRALFKAT 1810

RESULT 12

A48157  
 renal transcription factor Kid-1 - rat  
 A:Alternate names: zinc finger protein Kid-1  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 17-Mar-1999  
 C:Accession: A48157; A54304; I57401  
 R:Witzgall, R.; O'Leary, E.; Gessner, R.; Ouellette, A.J.; Bonventre, J.V.  
 Mol. Cell. Biol. 13, 1933-1942, 1993  
 A:Title: Kid-1, a putative renal transcription factor: regulation during ontogeny and in  
 A:Reference number: A48157; MUID:93180842  
 A:Accession: A48157  
 A:Molecule type: mRNA  
 A:Residues: 1-576 <WIT>  
 A:CROSS-references: GB:M96548; NID:g205066; PID:g205067  
 A:Experimental source: kidney  
 R:Witzgall, R.; Volk, R.; Yeung, R.S.; Bonventre, J.V.  
 Genomics 20, 203-209, 1994  
 A:Title: Genomic structure and chromosomal location of the rat gene encoding the zinc fi  
 A:Reference number: A54304; MUID:94292199  
 A:Accession: A54304  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-110 <WIT>  
 C:Keywords: transcription factor; zinc finger

Query Match 66.0%; Score 35; DB 2; Length 576;  
 Best Local Similarity 85.7%; Pred. No. 48;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLKFKCK 8

Db 213 EKLKFKCK 219

RESULT 13

C71634  
 serine hydroxymethyltransferase (glyA) RP743 - Rickettsia prowazekii  
 C:Species: Rickettsia prowazekii  
 C>Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 29-Jan-1999  
 C:Accession: C71634  
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U  
 Nature 396, 133-140, 1998  
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A:Reference number: A71630  
 A:Accession: C71634  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-420 <AND>  
 A:CROSS-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PID:el343016; PID:g3861272

A:Experimental source: strain Madrid E  
 C:Genetics:  
 A:Gene: glyA; RP743  
 C:Superfamily: glycine hydroxymethyltransferase

Query Match 64.2%; Score 34; DB 2; Length 420;  
 Best Local Similarity 85.7%; Pred. No. 56;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKLKFKCK 8

Db 83 KKLKFKCK 89

RESULT 14

F64373  
 hypothetical protein homolog MJ0590 - Methanococcus jannaschii  
 C:Species: Methanococcus jannaschii  
 C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1998  
 C:Accession: F64373  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak  
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,  
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc  
 A:Reference number: A64300; MUID:96337999  
 A:Accession: F64373  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-704 <BUL>  
 A:CROSS-references: GB:U67507; GB:L77117; NID:gl591288; PID:gl591298; TIGR:MJ0590; PI  
 C:Genetics:  
 A:Map position: FOR522030-524144  
 C:Superfamily: hypothetical protein MJ0590

Query Match 64.2%; Score 34; DB 2; Length 704;  
 Best Local Similarity 85.7%; Pred. No. 84;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFKCK 7

Db 521 CKKLGKCK 527

RESULT 15

G65110  
 hypothetical 35.2 kD protein in murZ-rpoN intergenic region - Escherichia coli (strai  
 C:Species: Escherichia coli  
 C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 29-Jan-1999  
 C:Accession: G65110  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
 .A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: G65110  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-328 <BLAT>  
 A:CROSS-references: GB:AE000399; GB:U00096; NID:g2367201; PID:gl789588; UWGP:b3197  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: yrbH  
 C:Superfamily: probable ATP-binding protein gutQ; CBS homology

Query Match 64.2%; Score 34; DB 2; Length 328;  
 Best Local Similarity 62.5%; Pred. No. 46;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKKLFCK 8  
|:|:|  
Db 42 CERMFCK 49

Search completed: September 7, 1999, 23:22:49  
Job time: 956 sec

**THIS PAGE BLANK (USPTO)**



Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	43	81.1	414	1	GLVA_CAMYE	P24531 campylobact
2	39	73.6	32	1	TX29_PHONE	P29426 phoneutria
3	39	73.6	803	1	ZN43_HUMAN	P38160 homo sapien
4	37	69.8	402	1	HEM1_METH	Q27093 methanobact
5	37	69.8	527	1	NU2M_ACACA	Q37376 acanthamoeb
6	36	67.9	407	1	YEAD_SCHPO	O14078 schizosacch
7	35	66.0	572	1	KID1_MOUSE	Q61751 mus musculu
8	35	66.0	576	1	KID1_RAT	Q02975 rattus norv
9	34.5	65.1	931	1	EMR1_MOUSE	Q61549 mus musculu
10	34	64.2	420	1	GLYA_RICPR	O08370 rickettsia
11	34	64.2	704	1	Y590_METYA	Q58010 methanococc
12	34	64.2	118	1	YB32_METYA	Q58532 methanococc
13	34	64.2	116	1	YEB0_YEAST	P40000 saccharomyc
14	34	64.2	204	1	YMO9_PARTE	P15610 paramecium
15	34	64.2	328	1	YRBH_ECOLI	P45395 escherichia
16	34	64.2	1191	1	ZN91_HUMAN	Q05481 homo sapien
17	33	62.3	239	1	DHSB_PORPU	R80477 porphyra pu
18	33	62.3	522	1	GAGS_HUMAN	P34059 homo sapien
19	33	62.3	397	1	PGIR_BRANA	P35337 brassica na
20	33	62.3	568	1	RPPP_PILHB	P32530 human parai
21	33	62.3	568	1	RPPP_PILHC	P28054 human parai
22	33	62.3	568	1	RPPP_PILHD	P32531 human parai
23	33	62.3	568	1	RPPP_PILHE	P32532 human parai
24	33	62.3	301	1	VU51_HSV6U	P52382 herpes simp
25	33	62.3	301	1	VU51_HSV6Z	P52542 herpes simp
26	33	62.3	371	1	YKOA_CAEEL	P34306 caenorhabdi
27	32	60.4	554	1	NBL4_MOUSE	P52963 mus musculu
28	32	60.4	691	1	NUAM_RECAM	O21241 reclinomona
29	32	60.4	889	1	RPAL_METYA	P41556 methanococc
30	32	60.4	377	1	SYW_SYNY3	P73655 synecthocyst
31	32	60.4	630	1	VF12_FOWPV	P36317 fowlpox vir
32	32	60.4	75	1	YE05_METYA	Q58800 methanococc
33	32	60.4	418	1	Z117_HUMAN	Q03924 homo sapien
34	32	60.4	168	1	Z642_XENLA	P18720 xenopus lae
35	32	60.4	163	1	ZN92_HUMAN	Q03936 homo sapien
36	31	58.5	422	1	5H1A_HUMAN	R08908 homo sapien
37	31	58.5	421	1	5H1A_MOUSE	Q64264 mus musculu
38	31	58.5	422	1	5H1A_RAT	P19327 rattus norv
39	31	58.5	150	1	CNRD_BOVIN	Q95142 bos taurus
40	31	58.5	1114	1	DEPI_CHICK	P36197 gallus gall
41	31	58.5	517	1	DOR_DROME	P49762 drosophila
42	31	58.5	1029	1	END1_YEAST	P12868 saccharomyc
43	31	58.5	82	1	ESCA_CALVI	Q01797 calliphora

OS PHONEUTRIA NIGRIVENTER (BRAZILIAN ARMED SPIDER).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; CHELICERATA; ARACHNIDA; ARANEAE;  
 OC ARANEOMORPHAE; ENTELEGYNAE; LYCOSOIDAE; CTENIDAE; PHONEUTRIA.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=VENOM;  
 RX MEDLINE: 93011905  
 RA CORDEIRO M.N., DINIZ C.R., VALENTIM A.C., VON EICKSTEDT V.R.D.,  
 RA GILROY J., RICHARDSON M.;  
 RT "The purification and amino acid sequences of four Tx2 neurotoxins  
 RT from the venom of the Brazilian 'armed' spider Phoneutria nigriventer  
 RT (Keys).";  
 RL FEBS LETT. 310:153-156(1992).  
 CC -!- FUNCTION: CAUSES TAIL ERECTION, SCRATCHING AND A REDUCTION IN  
 CC MOBILITY AT A DOSE LEVEL OF 1.40 MG/MOUSE.  
 CC PIR: S29217; S29217.  
 DR VENOM; NEUROTOXIN.  
 KW VENOM; NEUROTOXIN.  
 SQ SEQUENCE 32 AA; 3743 MW; 3E6E1AD0 CRC32;  
 Query Match 73.6%; Score 39; DB 1; Length 32;  
 Best Local Similarity 77.8%; Pred. No. 0.54;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CKLKFKCKT 9  
 DB 15 CCKKFKCKT 23  
 RESULT 3  
 ZN43 HUMAN STANDARD; PRT; 803 AA.  
 ID ZN43 HUMAN STANDARD; PRT; 803 AA.  
 AC P28160;  
 DT 01-DEC-1992 (REL. 24, CREATED)  
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE ZINC FINGER PROTEIN 43 (ZINC PROTEIN HTP6).  
 GN ZNF43.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=T-CELL;  
 RX MEDLINE: 91279444.  
 RA LOVERING R., TROMSDALE J.;  
 RT "A gene encoding 22 highly related zinc fingers is expressed in  
 RT lymphoid cell lines.";  
 RL NUCLEIC ACIDS RES. 19:2921-2927(1991).  
 RN [2]  
 RP SEQUENCE OF 38-190 FROM N.A.  
 RX MEDLINE: 91219421.  
 RA BELLEFROID E.J., PONCELET D.A., LECOCQ P.J., REVELANT O.,  
 RA MARTIAL J.A.;  
 RT "The evolutionarily conserved Kruppel-associated box domain defines a  
 RT subfamily of eukaryotic multifingered proteins";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:3608-3612(1991).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- TISSUE SPECIFICITY: T AND B CELL LINES.  
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -!- SIMILARITY: CONTAINS 1 KRAB BOX.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X59244; G38032; -;  
 DR EMBL; M61869; G184450; -;

DR PIR: S26823; S26823.  
 DR PIR: D39384; D39384.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 19.  
 DR PFAM: PF00096; 2f-C2H2; 21.  
 DR HSSP: P25490; 1ZNM.  
 KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; NUCLEAR PROTEIN.  
 FT DOMAIN 1 37  
 FT KRAB BOX ("A BOX").  
 FT KRAB BOX ("B BOX").  
 FT DOMAIN 38 69  
 FT 22 C2H2-TYPE ZINC-FINGERS.  
 FT DOMAIN 168 778  
 FT 22 C2H2-TYPE ZINC-FINGERS.  
 FT ZN\_FING 168 190  
 FT C2H2-TYPE.  
 FT ZN\_FING 196 218  
 FT C2H2-TYPE (DEGENERATE).  
 FT ZN\_FING 224 246  
 FT C2H2-TYPE (DEGENERATE).  
 FT ZN\_FING 252 274  
 FT C2H2-TYPE (DEGENERATE).  
 FT ZN\_FING 280 302  
 FT C2H2-TYPE.  
 FT ZN\_FING 308 330  
 FT C2H2-TYPE.  
 FT ZN\_FING 336 358  
 FT C2H2-TYPE.  
 FT ZN\_FING 364 386  
 FT C2H2-TYPE.  
 FT ZN\_FING 392 414  
 FT C2H2-TYPE.  
 FT ZN\_FING 420 442  
 FT C2H2-TYPE.  
 FT ZN\_FING 448 470  
 FT C2H2-TYPE.  
 FT ZN\_FING 476 498  
 FT C2H2-TYPE.  
 FT ZN\_FING 504 526  
 FT C2H2-TYPE.  
 FT ZN\_FING 532 554  
 FT C2H2-TYPE.  
 FT ZN\_FING 560 582  
 FT C2H2-TYPE.  
 FT ZN\_FING 588 610  
 FT C2H2-TYPE.  
 FT ZN\_FING 616 638  
 FT C2H2-TYPE.  
 FT ZN\_FING 644 666  
 FT C2H2-TYPE.  
 FT ZN\_FING 672 694  
 FT C2H2-TYPE.  
 FT ZN\_FING 700 722  
 FT C2H2-TYPE.  
 FT ZN\_FING 728 750  
 FT C2H2-TYPE.  
 FT ZN\_FING 756 778  
 FT C2H2-TYPE.  
 FT CONFLICT 181 181  
 FT P -> S (IN REF. 2).  
 SQ SEQUENCE 803 AA; 93487 MW; DB052880 CRC32;  
 Query Match 73.6%; Score 39; DB 1; Length 803;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KKLKFKCK 8  
 DB 165 KKLKFKCK 171  
 RESULT 4  
 HEMI METH STANDARD; PRT; 402 AA.  
 ID HEMI METH STANDARD; PRT; 402 AA.  
 AC O27093;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE GLUTAMYL-TRNA REDUCTASE (EC 1.2.1.-) (GLUTR).  
 GN HEMA OR MTH1012.  
 OS METHANOBACTERIUM THERMOAUTOTROPHICUM.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOBACTERIALES; METHANOBACTERIACEAE;  
 OC METHANOBACTERIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DELTA H.  
 RX MEDLINE: 98037514.  
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
 RA ALDRIDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,  
 RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTRIER B., QIU D.,  
 RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
 RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL J., PRABHAKAR S.,  
 RA MCDUGALL S., SHIMER G., GOVAL A., PIETROVSKI J., CHURCH G.M.,  
 RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics.";  
 RL J. BACTERIOL. 179:7135-7155(1997).  
 CC -!- CATALYTIC ACTIVITY: GLUTAMYL-TRNA (GLU) + NADPH = GLUTAMATE-1-  
 CC SEMIALDEHYDE + NADP(+) + TRNA (GLU).  
 CC -!- PATHWAY: FIRST STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.

CC -1- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE000874; G2622114; -  
 DR PROSITE: PS00747; GLUTR; 1.  
 DR PFAM: PF00745; GlucR; 1.  
 DR PORPHYRIN BIOSYNTHESIS: OXIDOREDUCTASE; NADP.  
 SQ SEQUENCE 402 AA; 44940 MW; 66A78E87 CRC32;

Query Match 69.8%; Score 37; DB 1; Length 402;  
 Best Local Similarity 75.0%; Pred. No. 9.5;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CKKLFCKK 8  
 DB 392 CAKLFCK 399  
 |||||

RESULT 5  
 NU2M\_ACACA STANDARD; PRT; 527 AA.  
 AC Q37376;  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).  
 GN ND2 OR NAD2  
 OS ACANTHAMOEBA CASTELLANII (AMOEBA).  
 OC EUKARYOTA; ACANTHAMOEBIIDAE; ACANTHAMOEBA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 30010 / NEFF;  
 RA BURGER G., PLANTE I., LONGERAN K.M., GRAY M.W.;  
 RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U12386; G562039; -  
 DR PFAM: PF00361; oxidored\_q1; 1.  
 DR OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION; TRANSMEMBRANE.  
 SQ SEQUENCE 527 AA; 61407 MW; D7FC10C8 CRC32;

Query Match 69.8%; Score 37; DB 1; Length 527;  
 Best Local Similarity 87.5%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 KKLKCKT 9  
 DB 13 KKLKCKGT 20  
 |||||

RESULT 6  
 YEAD\_SCHPO STANDARD; PRT; 407 AA.  
 ID YEAD\_SCHPO  
 AC O14078;  
 DT 15-JUL-1998 (REL. 36, CREATED)

DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 45.4 KD GTP-BINDING PROTEIN C2E11.13C IN CHROMOSOME I.  
 GN SPAC2E11.13C.  
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;  
 OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;  
 CC SCHIZOSACCHAROMYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WOOD V.;  
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SIMILARITY: BELONGS TO THE YCHF FAMILY OF GTP-BINDING PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL031180; E1314272; -  
 DR HYPOTHETICAL PROTEIN; GTP-BINDING.  
 KW NP\_BIND 52 59 GTP (POTENTIAL).  
 FT NP\_BIND 106 110 GTP (POTENTIAL).  
 SQ SEQUENCE 407 AA; 45362 MW; 877F862F CRC32;

Query Match 67.9%; Score 36; DB 1; Length 407;  
 Best Local Similarity 66.7%; Pred. No. 14;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CKKLFCKKT 9  
 DB 380 CKSVGKCKT 388  
 ||:||||

RESULT 7  
 KIDL\_MOUSE STANDARD; PRT; 572 AA.  
 ID KIDL\_MOUSE  
 AC Q61751;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17).  
 GN TCF17 OR KIDL.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LENS;  
 RA BRADY J.P., DUNCAN M.K., WAMROUSEK E.F., PIATIGORSKY J.;  
 RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- FUNCTION: IT MAY PLAY A ROLE IN RENAL DEVELOPMENT AND MAY  
 CC ALSO BE INVOLVED IN THE REPAIR OF THE KIDNEY AFTER ISCHEMIA-  
 CC REPERFUSION OR FOLIC ACID ADMINISTRATION.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 1 KRAB BOX.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L77247; G1256362; -

DR MGD: MGI:103172; TCF17.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 13.  
 DR PFAM: PF00096; zf-C2H2; 13.  
 DR HSP: P08047; ISP2.  
 KW ZINC-FINGER: TRANSCRIPTION REGULATION; DNA-BINDING; ACTIVATOR;  
 KW NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN; METAL-BINDING.  
 FT DOMAIN 12 53  
 FT DOMAIN 54 85  
 FT DOMAIN 184 567  
 FT ZN\_FING 184 206  
 FT ZN\_FING 212 234  
 FT ZN\_FING 240 262  
 FT ZN\_FING 268 290  
 FT ZN\_FING 321 343  
 FT ZN\_FING 349 371  
 FT ZN\_FING 377 399  
 FT ZN\_FING 405 427  
 FT ZN\_FING 431 455  
 FT ZN\_FING 461 483  
 FT ZN\_FING 489 511  
 FT ZN\_FING 517 539  
 FT ZN\_FING 545 567  
 SQ SEQUENCE 572 AA; 65720 MW; 4F7FBDC CRC32;

Query Match 66.0%; Score 35; DB 1; Length 572;  
 Best Local Similarity 85.7%; Pred. No. 28;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLFCCK 8  
 :|||||  
 Db 209 EKLFCK 215

## RESULT 8

ID KID1\_RAT STANDARD; PRT; 576 AA.  
 AC Q02975;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17).  
 GN TCF17 OR KID1.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-KIDNEY;  
 RX MEDLINE; 93180842.  
 RA WITZGALL R., O'LEARY E., GESSNER R., OUELLETTE A.J., BONVENTRE J.V.;  
 RT "Kid-1, a putative renal transcription factor: regulation during  
 RT ontogeny and in response to ischemia and toxic injury.";  
 RL MOL. CELL. BIOL. 13:1933-1942(1993).  
 RN [2]  
 RP SEQUENCE OF 1-110 FROM N.A.  
 RC TISSUE-LIVER;  
 RX MEDLINE; 94292199.  
 RA WITZGALL R., VOLK R., YEUNG R.S., BONVENTRE J.V.;  
 RT "Genomic structure and chromosomal location of the rat gene encoding  
 RT the zinc finger transcription factor Kid-1.";  
 RL GENOMICS 20:203-209(1994).  
 CC -1- FUNCTION: IT MAY PLAY A ROLE IN RENAL DEVELOPMENT AND MAY  
 CC ALSO BE INVOLVED IN THE REPAIR OF THE KIDNEY AFTER ISCHEMIA-  
 CC REPERFUSION OR POLIC ACID ADMINISTRATION.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE KIDNEY.  
 CC -1- DEVELOPMENTAL STAGE: PROTEIN LEVELS ACCUMULATE WITH AGE IN  
 CC POSTNATAL RENAL DEVELOPMENT.  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL SUBFAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 1 KRAB BOX.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M95548; G205067; -  
 EMBL; S73488; G688395; -  
 EMBL; S73484; G688395; JOINED.  
 EMBL; S73485; G688395; JOINED.  
 EMBL; S73486; G688395; JOINED.  
 PROSITE: PS00028; ZINC\_FINGER\_C2H2; 13.  
 PFAM: PF00096; zf-C2H2; 13.  
 HSP: P08047; ISP2.  
 TRANSFAC: T02326; -  
 ZINC-FINGER; TRANSCRIPTION REGULATION; DNA-BINDING; ACTIVATOR;  
 NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN; METAL-BINDING.

FT DOMAIN 12 54  
 FT DOMAIN 55 86  
 FT DOMAIN 96 111  
 FT DOMAIN 170 189  
 FT DOMAIN 190 552  
 FT ZN\_FING 190 210  
 FT ZN\_FING 218 238  
 FT ZN\_FING 246 266  
 FT ZN\_FING 274 294  
 FT ZN\_FING 308 328  
 FT ZN\_FING 336 356  
 FT ZN\_FING 364 384  
 FT ZN\_FING 392 412  
 FT ZN\_FING 420 440  
 FT ZN\_FING 448 468  
 FT ZN\_FING 476 496  
 FT ZN\_FING 504 524  
 FT ZN\_FING 532 552  
 SQ SEQUENCE 576 AA; 66253 MW; 6F583606 CRC32;

Query Match 66.0%; Score 35; DB 1; Length 576;  
 Best Local Similarity 85.7%; Pred. No. 28;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLFCCK 8  
 :|||||  
 Db 213 EKLFCK 219

## RESULT 9

ID EMR1\_MOUSE STANDARD; PRT; 931 AA.  
 AC Q61549;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE CELL SURFACE GLYCOPROTEIN EMR1 PRECURSOR (EMR1 HORMONE RECEPTOR)  
 DE (CELL SURFACE GLYCOPROTEIN F4/80).  
 GN EMR1 OR GPF480  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-PERITONEAL CAVITY;  
 RX MEDLINE; 96132946.  
 RA MCKNIGHT A.J., MACFARLANE A.J., DRI P., TURLEY L., WILLIS A.C.,  
 RA GORDON S.;  
 RT "Molecular cloning of F4/80, a murine macrophage-restricted cell  
 RT surface glycoprotein with homology to the G-protein-linked  
 RT transmembrane 7 hormone receptor family.";  
 RL J. BIOL. CHEM. 271:486-489(1996).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA MEDLINE; 97312684.  
 RA LIN H.H., STUBBS L.J., MUCENSKI M.L.;  
 RT "Identification and characterization of a seven transmembrane hormone  
 receptor using differential display.";  
 RL GENOMICS 41:301-308(1997)  
 CC -1- FUNCTION: PROBABLY INVOLVED IN CELL ADHESION WITHIN TISSUES  
 CC AND RECEPTOR SIGNALING.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: IN MACROPHAGES; BUT ABSENT FROM THOSE WHICH  
 CC ARE LOCALIZED WITHIN T-CELL AREAS OF LYMPH NODES AND SPLEEN.  
 CC LOW LEVEL OF EXPRESSION ON BLOOD MONOCYTES.  
 CC -1- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X93328; E214364; -;  
 DR EMBL; U66888; G2078504; -;  
 DR GCRDB; GCR\_1309; -;  
 DR MGD; MGI:105054; GPF480.  
 DR PROSITE; PS00649; G\_PROTEIN\_RECF\_F2\_1; FALSE\_NEG.  
 DR PROSITE; PS00650; G\_PROTEIN\_RECF\_F2\_2; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 6.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01187; EGF\_CA; 5.  
 DR PFAM; PF00008; EGF; 1.  
 DR HSP; P07204; 1FGD.  
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; RECEPTOR; GLYCOPROTEIN;  
 KW EGF-LIKE DOMAIN; REPEAT; SIGNAL.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 931 CELL SURFACE GLYCOPROTEIN F4/80.  
 FT DOMAIN 28 644 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 645 672 POTENTIAL.  
 FT DOMAIN 673 679 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 680 701 POTENTIAL.  
 FT DOMAIN 702 711 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 712 735 POTENTIAL.  
 FT DOMAIN 736 754 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 755 776 POTENTIAL.  
 FT DOMAIN 777 792 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 793 821 POTENTIAL.  
 FT DOMAIN 822 839 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 840 859 POTENTIAL.  
 FT DOMAIN 860 874 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 875 897 POTENTIAL.  
 FT DOMAIN 898 931 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 932 80 EGF-LIKE 1.  
 FT DOMAIN 81 132 EGF-LIKE 2.  
 FT DOMAIN 133 172 EGF-LIKE 3.  
 FT DOMAIN 173 221 EGF-LIKE 4.  
 FT DOMAIN 222 271 EGF-LIKE 5.  
 FT DOMAIN 272 318 EGF-LIKE 6.  
 FT DOMAIN 319 367 EGF-LIKE 7.  
 FT SITE 506 508 CELL ATTACHMENT SITE (POTENTIAL).  
 FT DISULFID 36 48 BY SIMILARITY.  
 FT DISULFID 42 57 BY SIMILARITY.  
 FT DISULFID 59 79 BY SIMILARITY.  
 FT DISULFID 85 98 BY SIMILARITY.  
 FT DISULFID 92 107 BY SIMILARITY.  
 FT DISULFID 109 131 BY SIMILARITY.  
 FT DISULFID 137 149 BY SIMILARITY.  
 FT DISULFID 143 158 BY SIMILARITY.  
 FT DISULFID 160 171 BY SIMILARITY.  
 FT DISULFID 177 189 BY SIMILARITY.  
 FT DISULFID 183 198 BY SIMILARITY.  
 FT DISULFID 200 220 BY SIMILARITY.

FT DISULFID 226 239 BY SIMILARITY.  
 FT DISULFID 233 248 BY SIMILARITY.  
 FT DISULFID 250 270 BY SIMILARITY.  
 FT DISULFID 276 286 BY SIMILARITY.  
 FT DISULFID 280 295 BY SIMILARITY.  
 FT DISULFID 297 317 BY SIMILARITY.  
 FT DISULFID 323 336 BY SIMILARITY.  
 FT DISULFID 330 345 BY SIMILARITY.  
 FT DISULFID 347 366 BY SIMILARITY.  
 FT CARBOHYD 148 148 POTENTIAL.  
 FT CARBOHYD 167 167 POTENTIAL.  
 FT CARBOHYD 229 229 POTENTIAL.  
 FT CARBOHYD 269 269 POTENTIAL.  
 FT CARBOHYD 283 283 POTENTIAL.  
 FT CARBOHYD 405 405 POTENTIAL.  
 FT CARBOHYD 417 417 POTENTIAL.  
 FT CARBOHYD 474 474 POTENTIAL.  
 FT CARBOHYD 498 498 POTENTIAL.  
 FT CARBOHYD 706 706 POTENTIAL.  
 SQ SEQUENCE 931 AA; 102129 MW; 35DB95D7 CRC32;  
 Query Match 65.18; Score 34.5; DB 1; Length 931;  
 Best Local Similarity 77.8%; Pred. No. 51;  
 Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 CKK-LFKCK 8  
 II: |||||  
 Db 366 CKRILFKCK 374  
 RESULT 10  
 GLYA\_RICPR  
 ID GLYA\_RICPR STANDARD; PRT; 420 AA.  
 AC 008370; 1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)  
 DE (SHMT).  
 GN GLYA OR RP743.  
 OS RICKETTSIA PROWAZEKII.  
 OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RICKETTSIALES;  
 OC RICKETTSIAEAE; RICKETTSIAEAE; RICKETTSIA.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-MADRID E;  
 RX MEDLINE; 97419517.  
 RA ANDERSSON J.O., ANDERSSON S.G.E.;  
 RT "Genomic rearrangements during evolution of the obligate  
 RT intracellular parasite Rickettsia prowazekii as inferred from an  
 RT analysis of 52015 bp nucleotide sequence.";  
 RT MICROBIOLOGY 143:2783-2795(1997).  
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +  
 CC H(2)O -> TETRAHYDROFOLATE + L-SERINE.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,  
 CC HORMONES AND OTHER COMPONENTS.  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y11778; E307841; -;  
 DR PROSITE; PS00096; SHMT; 1.

```

DR PFAM: PF00464; SHMT; 1.
KW TRANSFERASE; PYRIDOXAL PHOSPHATE; ONE-CARBON METABOLISM.
FT BINDING 230 230 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 420 AA; 46381 MW; 4C8ACBEE CRC32;

Query Match      64.28; Score 34; DB 1; Length 420;
Best Local Similarity 85.78; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KKLKFK 8
Db 83 KKLKFK 89

RESULT 11
Y590_METJA
ID Y590_METJA STANDARD; PRT; 704 AA.
AC Q58010;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0590.
GN MJ0590.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE UPF0058 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67507; GI591298; -.
CC TIGR: MJ0590; -.
CC KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 704 AA; 78172 MW; 148F4430 CRC32;

Query Match      64.28; Score 34; DB 1; Length 704;
Best Local Similarity 85.78; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CKKLKFK 7
Db 521 CKKLKGC 527

RESULT 12
YB32_METJA
ID YB32_METJA STANDARD; PRT; 118 AA.
AC Q58532;
DT 15-JUL-1998 (REL. 36, CREATED)

```

```

DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ1132.
GN MJ1132.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE UPF0058 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67556; GI499985; -.
CC TIGR: MJ1132; -.
CC KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 118 AA; 13841 MW; 35681C2D CRC32;

Query Match      64.28; Score 34; DB 1; Length 118;
Best Local Similarity 100.08; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CKKLKFK 6
Db 32 CKKLKFK 37

RESULT 13
YEB0_YEAST
ID YEB0_YEAST STANDARD; PRT; 116 AA.
AC P40000;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 13.5 KD PROTEIN IN GIC3-GC4 INTERGENIC REGION.
GN YEL010W.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMICETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-S288C / AB972;
RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,
RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,
RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNICKE-SMITH S.,
RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,
RA MOSEDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,
RA PETEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOENEN T., SMITH V.,
RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN D., DAVIS R.W.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U18530; G502377; -  
 CC HYPOTHETICAL PROTEIN.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 116 AA; 13523 MW; 593E478D CRC32;

Query Match 64.2%; Score 34; DB 1; Length 116;  
 Best Local Similarity 50.0%; Pred. No. 12;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCK 8  
 |::|::|  
 Db 70 CRLYECK 77

## RESULT 14

YMO9\_PARTE STANDARD; PRT; 204 AA.  
 AC P15610;  
 DT 01-APR-1990 (REL. 14, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 22.9 KD PROTEIN (ORF9).  
 OS PARAMETICUM TETRAURELIA.  
 OG MITOCHONDRION.  
 OC EUKARYOTA; ALVEOLATA; CILIOPHORA; NASSOPHOREA; PENICULIDA; PARAMETICUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-STOCK 51;  
 RX MEDLINE: 90174913.  
 RA PRITCHARD A.E., SEITZHAMER J.J., MAHALINGAM R., SABLE C.L.,  
 RA VENUTI S.E., CUMMINGS D.J.;  
 RT "Nucleotide sequence of the mitochondrial genome of Parametrium.";  
 RL NUCLEIC ACIDS RES. 18:173-180(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X15917; G578758; -  
 CC PIR: S07741; S07741.  
 KW HYPOTHETICAL PROTEIN; MITOCHONDRION.  
 SQ SEQUENCE 204 AA; 22919 MW; 27F9BFCF CRC32;

Query Match 64.2%; Score 34; DB 1; Length 204;  
 Best Local Similarity 75.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCK 8  
 |||||  
 Db 114 CKRFFK 121

## RESULT 15

YRBH\_ECOLI STANDARD; PRT; 328 AA.  
 AC P45395;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 35.2 KD PROTEIN IN MURA-RPON INTERGENIC REGION (O328).  
 OS YRBH.  
 OS ESCHERICHIA COLI.

OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 OC ESCHERICHIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE: 97426617  
 RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,  
 RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,  
 RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,  
 RA MAU B., SHAO Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL SCIENCE 277:1453-1474(1997)  
 CC -1- SIMILARITY: STRONG, TO H. INFLUENZAE H11678.  
 CC -1- SIMILARITY: TO E. COLI GUTQ AND KPSE.  
 CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U18997; G606135; -  
 CC EMBL: AE000399; G1789588; -  
 CC ECOGENE; EG12803; YRBH.  
 DR PFAM: PF00571; CBS: 2.  
 KW HYPOTHETICAL PROTEIN; ATP-BINDING; REPEAT; CBS DOMAIN.  
 FT NP\_BIND 56 61 ATP (POTENTIAL).  
 FT DOMAIN 208 263 CBS 1.  
 FT DOMAIN 275 327 CBS 1.  
 SQ SEQUENCE 328 AA; 35196 MW; E7507FCA CRC32;

Query Match 64.2%; Score 34; DB 1; Length 328;  
 Best Local Similarity 62.5%; Pred. No. 27;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCK 8  
 |::|::|  
 Db 42 CERMFCK 49

Search completed: September 7, 1999, 23:59:15  
 Job time: 527 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:45 ; Search time 160.2 Seconds  
(without alignments)  
3.458 Million cell updates/sec

Title: US-09-124-280A-34

Perfect score: 53

Sequence: 1 CKKLFCKT 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- SPTREMBL10.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	75.5	1914	Q13129	Q13129 homo sapien
2	38	71.7	241	O67793	O67793 aquifex aeo
3	38	71.7	557	O43340	O43340 homo sapien
4	37	69.8	318	O56914	O56914 yersinia en
5	37	69.8	1077	O74853	O74853 schizosacch
6	37	69.8	1051	O26055	O26055 paracentrot
7	37	69.8	774	O18387	O18387 drosophila
8	37	69.8	3198	O26639	O26639 strongyloce
9	36	67.9	200	O94993	O94993 drosophila
10	35	66.0	3703	O15911	O15911 homo sapien
11	35	66.0	605	O60765	O60765 homo sapien
12	35	66.0	2783	O13719	O13719 homo sapien
13	35	66.0	146	O22087	O22087 petunia hyb
14	35	66.0	163	O22088	O22088 petunia hyb
15	35	66.0	311	O92066	O92066 arabidopsis
16	35	66.0	402	O92065	O92065 arabidopsis
17	35	66.0	3726	O61329	O61329 mus musculu
18	35	66.0	999	O20851	O20851 rice yellow
19	35	66.0	999	O20852	O20852 rice yellow
20	34.5	65.1	297	O08743	O08743 mus musculu
21	34.5	65.1	304	O08744	O08744 mus musculu
22	34.5	65.1	255	O08745	O08745 mus musculu
23	34	64.2	770	O06833	O06833 saccharomyc
24	34	64.2	393	O75260	O75260 homo sapien
25	34	64.2	501	O75261	O75261 homo sapien
26	34	64.2	314	O25583	O25583 onchocerca
27	34	64.2	447	O25635	O25635 onchocerca
28	34	64.2	263	O08128	O08128 onchocerca
29	34	64.2	86	O28774	O28774 pongo pygma

30	34	64.2	63	11	P97662	P97662 rattus norv
31	33	62.3	1087	4	O43279	O43279 homo sapien
32	33	62.3	683	4	O43387	O43387 homo sapien
33	33	62.3	367	4	O13398	O13398 homo sapien
34	33	62.3	683	4	O14978	O14978 homo sapien
35	33	62.3	353	4	O75437	O75437 homo sapien
36	33	62.3	1087	4	O99639	O99639 homo sapien
37	33	62.3	1087	4	O00267	O00267 homo sapien
38	33	62.3	135	5	O61729	O61729 ascaridia g
39	33	62.3	1317	5	O61855	O61855 caenorhabdi
40	33	62.3	675	5	O19377	O19377 caenorhabdi
41	33	62.3	391	10	O65905	O65905 arabidopsis
42	33	62.3	172	10	O22089	O22089 petunia hyb
43	33	62.3	399	10	O80603	O80603 arabidopsis
44	33	62.3	566	11	O62977	O62977 rattus norv
45	32	60.4	149	2	O87179	O87179 streptococc

## ALIGNMENTS

RESULT 1					
Q13129 PRELIMINARY; PRT; 1914 AA.					
ID	Q13129	AC	Q13129	DT	01-NOV-1996 (Tremblrel. 01, Created)
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)	DT	01-MAY-1999 (Tremblrel. 10, Last annotation update)	DE	ZN-15 RELATED ZINC FINGER PROTEIN (RLF).
GN	RLF.	OS	Homo sapiens (Human).	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eutheria; Primates; Catarrhini; Homnidae; Homo.	RC	[1]	RP	SEQUENCE FROM N.A.
RX	MEDLINE; 96132723.	RA	MAKEHA T.P., HELLSTEN E., VESA J., HIRVONEN H., PALOTIE A.,	RT	"The rearranged L-myc fusion gene (RLF) encodes a Zn-15 related zinc
RL	Oncogene 11:2699-2704(1995).	DR	EMBL; U23277; AAC50396.1; -.	DR	PFAM; PF000096; zf-C2H2; 14.
DR	PROSITE; P500028; ZINC_FINGER_C2H2; 14.	KW	zinc-finger; Metal-binding; DNA-binding.	SQ	SEQUENCE 1914 AA; 217898 MW; 05EBE74D CRC32;

Query Match 75.5%; Score 40; DB 4; Length 1914;  
Best Local Similarity 75.0%; Pred. NO. 9.9;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1 CKKLFCKK 8
DB	1358 CKKIFACK 1365

RESULT 2					
O67793 PRELIMINARY; PRT; 241 AA.					
ID	O67793	AC	O67793	DT	01-AUG-1998 (Tremblrel. 07, Created)
DT	01-AUG-1998 (Tremblrel. 07, Last sequence update)	DT	01-NOV-1998 (Tremblrel. 08, Last annotation update)	DE	HYPOPHETICAL 27.5 KD PROTEIN.
GN	AQ1986.	OS	Aquifex aeolicus.	OC	Bacteria; Aquificales; Aquificaceae; Aquifex.
OC	[1]	RP	SEQUENCE FROM N.A.	RC	STRAIN=VF5;
RX	MEDLINE; 98196666.	RA	DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,		

RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RL aeolicus.";  
 RL Nature 392:353-358 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR ENBL: AE00766; AAC0762.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 241 AA; 27514 MW; A94A054B CRC32;

Query Match 71.7%; Score 38; DB 2; Length 241;  
 Best Local Similarity 75.0%; Pred. No. 4.8;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCK 8  
 : : : : :  
 Db 162 CEKLLCK 169

RESULT 3  
 ID 043340 PRELIMINARY; PRT; 557 AA.  
 AC 043340;  
 DT 01-JUN-1998 (TRENBLrel. 06, Created)  
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE R28830.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LAMERDIN J.E., MCCREARY P.M., ADAMSON A.W., BURKHART-SCHULTZ K.,  
 RA GARCIA E., KYLE A., RAMIREZ M., STILWAGEN S., GARNES J., DANGANAN L.,  
 RA BRUCE R., QUAN G., MONTGOMERY M., OW D., KOBAYASHI A., OLSEN A.O.,  
 RA CARRANO A.V.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR ENBL: AC003682; AAC24609.1;  
 DR PFAM: PF00096; zf-C2H2; 13.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 13.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 557 AA; 63861 MW; ED322DFA CRC32;

Query Match 71.7%; Score 38; DB 4; Length 557;  
 Best Local Similarity 66.7%; Pred. No. 9.1;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCK 9  
 : : : : :  
 Db 426 CEKSFCK 434

RESULT 4  
 ID Q56914 PRELIMINARY; PRT; 318 AA.  
 AC Q56914;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE TRSB.  
 GN TRSB.  
 OS Versinia enterocolitica.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Versinia.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=6471776 (SEROTYPE O:3);  
 RX MEDLINE; 96100456.  
 RA SKURNIK M., VENHO R., TOIVANEN P., AL HENDY A.;  
 RT "A novel locus of Versinia enterocolitica serotype O:3 involved in  
 RT lipopolysaccharide outer core biosynthesis.";  
 RL Mol. Microbiol. 17:575-594 (1995).  
 DR ENBL: 247767; CAA87699.1;  
 DR PFAM: PF00535; Glycos\_transf.2; 1.  
 SQ SEQUENCE 318 AA; 37112 MW; 7ADB7E51 CRC32;

Query Match 69.8%; Score 37; DB 2; Length 318;  
 Best Local Similarity 62.5%; Pred. No. 9.1;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFCK 8  
 : : : : :  
 Db 193 CKSFECK 200

RESULT 5  
 ID 074853 PRELIMINARY; PRT; 1077 AA.  
 AC 074853;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE PUTATIVE CYSTINE-RICH TRANSCRIPTIONAL REGULATOR.  
 GN SPC18.03.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972H-;  
 RA HILBERT H., DUESTERHOFF A., WOOD V., RAJANDREAM M.A., BARRELL B.G.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR ENBL: AL031907; CAA21417.1;  
 DR PFAM: PF01424; R3H; 1.  
 DR PFAM: PF01422; zf-NF-X1; 8.  
 SQ SEQUENCE 1077 AA; 121087 MW; 99CE9378 CRC32;

Query Match 69.8%; Score 37; DB 3; Length 1077;  
 Best Local Similarity 66.7%; Pred. No. 23;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKLFCK 9  
 : : : : :  
 Db 702 CNKLSCK 710

RESULT 6  
 ID Q26055 PRELIMINARY; PRT; 1051 AA.  
 AC Q26055;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE 2-ALPHA COLLAGEN (COLL2-ALPHA) (FRAGMENT).  
 OS Paracentrotus lividus (Common sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea;  
 OC Euechinoidea; Echinacea; Echinoidae; Echinidae; Paracentrotus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90216744.  
 RA D'ALESSIO M., RAMIREZ F., SUZUKI H.R., SOLURSH M., GAMBINO R.;  
 RT "Cloning of a fibrillar collagen gene expressed in the mesenchymal  
 RT cells of the developing sea urchin embryo.";  
 RL J. Biol. Chem. 265:7050-7054 (1990).  
 DR ENBL: J05422; AAA29440.1;

DR PFAM: PF01410; COLFI; 1.  
 DR PFAM: PF01391; Collagen; 12.  
 FT NON\_TER 1  
 FT CHAIN 1 786 POTENTIAL.  
 SQ SEQUENCE 1051 AA; 101961 MW; 11D17FBD CRC32;

Query Match 69.8%; Score 37; DB 5; Length 1051;  
 Best Local Similarity 85.7%; Pred. No. 23;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFKC 7

DB 854 CKDLFKC 860

RESULT 7  
 O18387 PRELIMINARY; PRT; 774 AA.  
 AC Q18387;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE D19B.  
 GN D19B.

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
 OC Drosophilidae; Drosophila.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA JULLIEN D., KAS E.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ000261; CAA03977.1; -  
 DR FLYBASE: FBgn0022699; D19B.  
 DR PFAM: PF00096; zf-C2H2; 12.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 10.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 774 AA; 88059 MW; D52BE4AF CRC32;

Query Match 69.8%; Score 37; DB 5; Length 774;  
 Best Local Similarity 55.8%; Pred. No. 18;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFKCCT 9

DB 352 CEKVFRCPT 360

RESULT 8  
 Q26639 PRELIMINARY; PRT; 3198 AA.  
 AC Q26639;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE ALPHA-2 COLLAGEN.  
 GN COL2ALPHA.

OS Strongylocentrotus purpuratus (Purple sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea;  
 OC Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;  
 OC Strongylocentrotus.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 92348411.

RA EXPOSITO J.Y., D'ALESSIO M., SOLURSH M., RAMIREZ F.;  
 RT "Sea urchin collagen evolutionarily homologous to vertebrate  
 pro-alpha 2(I) collagen."  
 RL J. Biol. Chem. 267:15559-15562(1992).

RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE: 92381062.

RA EXPOSITO J.Y., D'ALESSIO M., RAMIREZ F.;  
 RT "Novel amino-terminal propeptide configuration in a fibrillar  
 RT procollagen undergoing alternative splicing."  
 RL J. Biol. Chem. 267:17404-17408(1992).  
 DR EMBL: M92041; AAA30040.1; -  
 DR PFAM: PF01410; COLFI; 1; -  
 DR PFAM: PF01391; Collagen; 17.  
 DR PFAM: PF00093; vvc; 1.  
 SQ SEQUENCE 3198 AA; 331465 MW; 9EF47D37 CRC32;

Query Match 69.8%; Score 37; DB 5; Length 3198;  
 Best Local Similarity 85.7%; Pred. No. 53;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKLFKC 7

DB 3002 CKDLFKC 3008

RESULT 9  
 Q94993 PRELIMINARY; PRT; 200 AA.  
 AC Q94993;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE RIBOSOMAL PROTEIN S6 GENE AND TWO POTENTIAL ALTERNATIVELY SPLICED  
 DE PROTEINS, COMPLETE CDS.  
 GN RPS6.

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
 OC Drosophilidae; Drosophila.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 93204998.

RA STEWART M.J., DENELL R.;  
 RT "Mutations in the Drosophila gene encoding ribosomal protein S6 cause  
 RT tissue overgrowth."

RL Mol. Cell. Biol. 13:2524-2535(1993).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94018587.

RA STEWART M.J., DENELL R.;  
 RT "The Drosophila ribosomal protein S6 gene includes a 3' triplication  
 RT that arose by unequal crossing-over."

RL Mol. Biol. Evol. 10:1041-1047(1993).  
 RN [3]

RP SEQUENCE FROM N.A.  
 RA STEWART M.J.;

RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L02074; AAB05983.1; -  
 DR EMBL: L02074; AAB05984.1; -  
 DR FLYBASE: FBgn0004922; Rps6.  
 DR PFAM: PF01092; Ribosomal\_S6e; 1.  
 DR PROSITE: PS00578; RIBOSOMAL\_S6e; 1.  
 SQ SEQUENCE 200 AA; 23072 MW; DC870C3B CRC32;

Query Match 67.9%; Score 36; DB 5; Length 200;  
 Best Local Similarity 66.7%; Pred. No. 9.8;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFKCCT 9

DB 87 CNKVRKCKT 95

RESULT 10

Q15911

ID Q15911

AC Q15911;

PRELIMINARY; PRT; 3703 AA.

DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE ZINC FINGER HOMEODOMAIN PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG;  
 RX MEDLINE; 92049333.  
 RA MORINAGA T., YASUDA H., HASHIMOTO T., HIGASHIO K., TAMAOKI T.;  
 RT "A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains  
 RT four homeodomains and seventeen zinc fingers.";  
 RL Mol. Cell. Biol. 11:6041-6049(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG;  
 RX MEDLINE; 96070776.  
 RA MIURA Y., TAM T., IDO A., MORINAGA T., MIKI T., HASHIMOTO T.,  
 RT "Cloning and characterization of an ATBF1 isoform that expresses in a  
 RT neuronal differentiation-dependent manner.";  
 RL J. Biol. Chem. 270:26840-26848(1995).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; L32832; AAC14462.1; -.  
 DR PFAM; PF00046; homeobox; 4.  
 DR PFAM; PF00096; zf-C2H2; 20.  
 DR PROSITE; PS00027; HOMEBOX\_1; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 14.  
 DR Homeobox; DNA-binding; Nuclear protein; Zinc-finger; Metal-binding.  
 KW VARIANT 3528 3529 SY -> AA.  
 FT VARIANT 3528 3529  
 SQ SEQUENCE 3703 AA; 404468 MW; 797F7644 CRC32;

Query Match 66.0%; Score 35; DB 4; Length 3703;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKCT 9  
 I: |||||  
 Db 2715 CRALEFCKT 2724

RESULT 11  
 O60765 PRELIMINARY; PRT; 605 AA.  
 ID O60765  
 AC O60765;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE HKL1.  
 GN HKL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FETAL BRAIN;  
 RX MEDLINE; 98127052.  
 RA OMORI Y., KYUSHIKI H., TAKEDA S., SUZUKI M., KAWAI A., FUJIWARA T.,  
 RA TAKAHASHI E., NAKAMURA Y.;  
 RT "Cloning, expression and mapping of a novel human zinc-finger gene  
 RT TCF17 homologous to rodent Kidx1.";  
 RL Cytogenet. Cell Genet. 78:285-288(1997).  
 DR EMBL; D89928; BAA25182.1; -.  
 DR PFAM; PF01352; KRAB; 1.  
 DR PFAM; PF00096; zf-C2H2; 13.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 13.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 605 AA; 69252 MW; B870EF8F CRC32;

Query Match 66.0%; Score 35; DB 4; Length 605;  
 Best Local Similarity 85.7%; Pred. No. 35;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLKFKCK 8  
 I: |||||  
 Db 239 EKLKFKCK 245

RESULT 12  
 Q13719 PRELIMINARY; PRT; 2783 AA.  
 ID Q13719  
 AC Q13719;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE ALPHA-FETOPROTEIN ENHANCER BINDING PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEPATOMA;  
 RX MEDLINE; 92049333.  
 RA MORINAGA T., YASUDA H., HIGASHIO K., TAMAOKI T.;  
 RT "A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains  
 RT four homeodomains and seventeen zinc fingers.";  
 RL Mol. Cell. Biol. 11:6041-6049(1991).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; D10250; BAA01095.1; -.  
 DR PFAM; PF00046; homeobox; 4.  
 DR PFAM; PF00096; zf-C2H2; 16.  
 DR PROSITE; PS00027; HOMEBOX\_1; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 11.  
 DR Homeobox; DNA-binding; Nuclear protein; Zinc-finger; Metal-binding.  
 KW VARIANT 2460 2460 A -> V.  
 FT VARIANT 2460 2460  
 SQ SEQUENCE 2783 AA; 305737 MW; DC7AAE23 CRC32;

Query Match 66.0%; Score 35; DB 4; Length 2783;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKCT 9  
 I: |||||  
 Db 1802 CRALEFCKT 1810

RESULT 13  
 O22087 PRELIMINARY; PRT; 146 AA.  
 ID O22087  
 AC O22087;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE ZPT2-7.  
 OS Petunia hybrida (Petunia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eukaryotes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Petunia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MITCHELL DIPLOID; TISSUE=ROOT;  
 RX MEDLINE; 98083196.  
 RA KUBO K., SAKAMOTO A., KOBAYASHI A., RYBKA Z., KANNO Y., NAKAGAWA H.,  
 RA NISHINO T., TAKATSUJI H.;  
 RT "Cys2/His2 zinc-finger protein family of petunia: evolution and  
 RT general mechanism of target-sequence recognition.";  
 RL Nucleic Acids Res. 26:608-615(1998).  
 DR EMBL; AB006602; BAA21924.1; -.  
 DR PFAM; PF00096; zf-C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 2.

KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 146 AA; 16517 MW; AF36F983 CRC32;

Query Match 66.0%; Score 35; DB 10; Length 146;  
 Best Local Similarity 62.5%; Pred. No. 12;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLFKCKT 9  
 :|:|:|  
 Db 34 RKIFECKT 41

## RESULT 14

O22088 PRELIMINARY; PRT; 163 AA.  
 AC O22088;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE ZPT2-8.  
 OS Petunia hybrida (Petunia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Petunia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MITCHELL DIPLOID; TISSUE=ROOT;  
 RX MEDLINE; 980833196.  
 RA KUBO K., SAKAMOTO A., KOBAYASHI A., RYBKA Z., KANNO Y., NAKAGAWA H.,  
 RA NISHINO T., TAKATSUJI H.;  
 RT "Cys2/His2 zinc-finger protein family of petunia: evolution and  
 RT general mechanism of target-sequence recognition.";  
 RL Nucleic Acids Res. 26:608-615(1998).  
 DR EMBL; AB006603; BAA21925.1; .  
 DR PFAM; PF00096; zf-C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 2.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 163 AA; 18513 MW; 9FD13F2B CRC32;

Query Match 66.0%; Score 35; DB 10; Length 163;  
 Best Local Similarity 62.5%; Pred. No. 13;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKLFKCKT 9  
 :|:|:|  
 Db 34 RKIFECKT 41

## RESULT 15

O9ZOG6 PRELIMINARY; PRT; 311 AA.  
 AC O9ZOG6;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE PUTATIVE POLYGALACTURONASE, 5' PARTIAL (FRAGMENT).  
 GN F9013.1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 OC eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA LIN X., KAUL S., SHEA T.P., FUJII C.Y., SHEN M., VANAKEN S.E.,  
 RA BARNSTEAD M.E., MASON T.M., BOWMAN C.L., RONNING C.M., BENITO M.,  
 RA CARRERA A.J., CREAMY T.H., BUELL C.R., TOWN C.D., NIERMAN W.C.,  
 RA FRASER C.M., VENTER J.C.;  
 RT "Arabidopsis thaliana chromosome II BAC F9013 genomic sequence.";

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES  
 CC (POLYGALACTURONASES).  
 DR EMBL; AC006248; AAD17390.1; .  
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.  
 KW Hydrolase; Glycosidase; Cell wall; Signal.  
 FT NON\_TER 1  
 SQ SEQUENCE 311 AA; 33904 MW; C274B280 CRC32;

Query Match 66.0%; Score 35; DB 10; Length 311;  
 Best Local Similarity 62.5%; Pred. No. 21;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKK 8  
 |:|:|  
 Db 265 CSKIFPCK 272

Search completed: September 7, 1999, 22:47:46  
 Job time: 7966 sec

**THIS PAGE BLANK (USPTO)**



PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;

DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 8; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00032;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11  
 |||||  
 Db 1 IKTKCKFLKCC 11

RESULT 3

ID W21623 standard; peptide; 11 AA.

AC W21623;

DT 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #35.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.

OS Synthetic.

FH Key Location/Qualifiers

FT disulfide bond 5..11

PN W09638163-A1.

PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;

DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic

PT amino acid sequence - reduces dose of antibiotic required

PS Claim 39; Page 28; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to

CC potentiate the activity of an antibiotic when they are co-administered

CC with the antibiotic. Compositions containing these peptides are used

CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of

CC the outer bacterial membrane to the antibiotic, allowing a reduction

CC in the dose of antibiotic required by 10-90% of the normal dose for

CC in vivo or in vitro application. Any toxic side effects are

CC correspondingly reduced.

CC Sequence 11 AA;

Query Match 100.0%; Score 62; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00032;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11  
 |||||  
 Db 1 IKTKCKFLKCC 11

RESULT 4

R33526

ID R33526 standard; peptide; 10 AA.

AC R33526;

DT 07-JUL-1993 (first entry)

DE Peptide for treating septic shock.

KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;

KW cytokine release control; treatment; pertussis; bacterial meningitis;

KW HIV related infections; polymyxin B.

OS Synthetic.

FH Key Location/Qualifiers

FT disulfide bond 4..10

PN ZA9200943-A.

PD 25-NOV-1992.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORRO M.

PI Porro M;

DR WPI; 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 4; Page 32; 39pp; English.

CC This is a specific example of a generic peptide of formula

CC R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where n = 1-100

CC and each R is H, an amino acid residue or a fatty acid residue.

CC The peptide is useful for treating or preventing septic shock.

CC mixing with polymyxin B to reduce its toxicity; removing

CC endotoxins from blood, sera or other fluids (in vivo or in

CC vitro); controlling release of cytokines induced by endotoxins;

CC as diagnostic reagents to detect and quantify toxins in blood

CC or sera; preparing non-toxic antigenic complexes of lipid A or

CC lipopolysaccharide (LPS); and for treating pertussis, bacterial

CC meningitis and HIV-related infections. The usual dose is 10-100

CC ug/kg/day, given parenterally. It binds to the same sites as

CC polymyxin B i.e. it inhibits all the toxic effects of lipid A. It

CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable

CC against proteases.

CC Sequence 10 AA;

Query Match 93.5%; Score 58; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
 |||||

Db 1 KTKCKFLKCC 10

RESULT 5

R39288

ID R39288 standard; peptide; 10 AA.

AC R39288;

DT 22-DEC-1993 (first entry)

DE Endotoxin lipid A neutralising peptide.

KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;

KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;

KW removal; sera; vaccines; drug solns.; solutions; cytokine release;

KW control; in vivo; in vitro; detoxification; detection;

OS Synthetic.

FH Key Location/Qualifiers

FT disulfide bond 4..10

PN W09314115-A.

PD 22-JUL-1993.

PF 14-MAY-1992; E01060.

PR 16-JAN-1992; US-819893.

PA (PORR/) PORRO M.

PI Porro M;



DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 forming non-toxic, antigenic complex, used to treat or prevent  
 septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 4; Page 32; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 component of bacterial endotoxin at the same site as polymyxin B  
 (PMB) and with about the same affinity to produce a non-toxic  
 antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 proteolytic degradation in serum, has no antibiotic activity and no  
 haemolytic action. It is especially used to treat or prevent septic  
 shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 sera, vaccines, drug solutions, etc.; to control release of cytokines  
 induced by endotoxins; for in vivo or in vitro detoxification of  
 bacterial endotoxins, and to detect or quantify endotoxins in blood  
 products.  
 SQ Sequence 10 AA;

Query Match 93.5%; Score 58; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
 |||||  
 DB 1 KTKCKFLKCC 10

## RESULT 6

W21619  
 ID W21619 standard; peptide; 10 AA.

AC W21619;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #31.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT Disulfide bond 4..10

PN WO9638163-A1.  
 PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;

DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
 amino acid sequence, reduces dose of antibiotic required

PS Claim 35; Page 27; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to  
 potentiate the activity of an antibiotic when they are co-administered  
 with the antibiotic. Compositions containing these peptides are used  
 to treat or prevent microbial infections. These peptides bind to  
 lipopolysaccharide on the bacteria so may increase permeability of  
 the outer bacterial membrane to the antibiotic, allowing a reduction  
 in the dose of antibiotic required by 10-90% of the normal dose for  
 in vivo or in vitro application. Any toxic side effects are  
 correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 93.5%; Score 58; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
 |||||  
 DB 1 KTKCKFLKCC 10

## RESULT 7

W62434  
 ID W62434 standard; peptide; 10 AA.

AC W62434;  
 DT 01-OCT-1998 (first entry)  
 DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:49.  
 KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;  
 KW bactericidal permeability increasing factor; B/Pi; neutralising;  
 KW human neutrophil granule bactericidal protein.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN US5786324-A.  
 PD 28-JUL-1998.  
 PF 24-MAR-1994; 218026.  
 PR 24-MAR-1994; US-218026.  
 PA (MINU) UNIV MINNESOTA.  
 PI Gray B, Haseman JR, Mayo K;  
 DR WPI; 98-436578/37.  
 PT Bactericidal and endotoxin-neutralising peptides - used in treating  
 e.g. Pseudomonas species infection and in protectively coating  
 prosthetic devices

PS Example 1; Column 18; 46pp; English.

CC The present invention describes bactericidal peptides (BP): (a) for  
 Pseudomonas species; (b) with endotoxin neutralising activity; (c)  
 with both endotoxin neutralising activity and bactericidal activity for  
 Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising  
 activity and bactericidal activity, the peptide having at least 1 amino  
 acid sequence of a peptide with endotoxin neutralising activity combined  
 with at least 1 different amino acid sequence of a peptide that has  
 bactericidal activity; and (e) a prosthetic device which has a  
 sufficient amount of BP attached to the surface to inhibit bacterial  
 growth. The peptides of the invention are used in treating bacterial  
 infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7 to  
 10-9 M, and Escherichia coli. The peptides are also used to treat  
 endotoxin shock. The present sequence represents a peptide derived  
 from human neutrophil granule bactericidal protein from an example of  
 the present invention.  
 SQ Sequence 10 AA;

Query Match 93.5%; Score 58; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
 |||||  
 DB 1 KTKCKFLKCC 10

## RESULT 8

W62435  
 ID W62435 standard; peptide; 11 AA.

AC W62435;  
 DT 01-OCT-1998 (first entry)  
 DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:50.  
 KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;  
 KW bactericidal permeability increasing factor; B/Pi; neutralising;  
 KW human neutrophil granule bactericidal protein.

OS Homo sapiens.

OS Synthetic.

PN US5786324-A.

PD 28-JUL-1998.

PF 24-MAR-1994; 218026.

PR 24-MAR-1994; US-218026.

PA (MINU) UNIV MINNESOTA.

PI Gray B, Haseman JR, Mayo K;

DR WPI; 98-436578/37.

PT Bactericidal and endotoxin-neutralising peptides - used in treating  
 e.g. Pseudomonas species infection and in protectively coating  
 prosthetic devices

PS Example 1; Column 18; 46pp; English.

CC The present invention describes bactericidal peptides (BP): (a) for  
 Pseudomonas species; (b) with endotoxin neutralising activity; (c)  
 with both endotoxin neutralising activity and bactericidal activity for  
 Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising  
 activity and bactericidal activity, the peptide having at least 1 amino

CC acid sequence of a peptide with endotoxin neutralising activity combined  
 CC with at least 1 different amino acid sequence of a peptide that has  
 CC bactericidal activity; and (e) a prosthetic device which has a  
 CC sufficient amount of BP attached to the surface to inhibit bacterial  
 CC growth. The peptides of the invention are used in treating bacterial  
 CC infection such as *Pseudomonas* strains e.g. for *P. aeruginosa* at 10-7 to  
 CC 10-9 M, and *Escherichia coli*. The peptides are also used to treat  
 CC endotoxin shock. The present sequence represents a peptide derived  
 CC from human neutrophil granule bactericidal protein from an example of  
 CC the present invention.  
 SQ Sequence 11 AA;

Query Match 93.5%; Score 58; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0014;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KTKCKFLKCC 11  
 |||||  
 DB 2 KTKCKFLKCC 11

RESULT 9  
 R33525  
 ID R33525 standard; peptide; 7 AA.  
 AC R33525;  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1. .7  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 3; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)<sub>n</sub>-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood;  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 7 AA;

Query Match 69.4%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 CKFLKCC 11  
 |||||  
 DB 1 CKFLKCC 7

RESULT 10  
 R39287  
 ID R39287 standard; peptide; 7 AA.  
 AC R39287; 1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW Polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1. .7  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 3; Page 32; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 7 AA;

Query Match 69.4%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 CKFLKCC 11  
 |||||  
 DB 1 CKFLKCC 7

RESULT 11  
 W21618  
 ID W21618 standard; peptide; 7 AA.  
 AC W21618; 1997 (first entry)  
 DE Antibiotic potentiating peptide #30.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1. .7  
 PN WO9638163-A.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 34; Page 27; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 7 AA;

Query Match 69.4%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CKFLKCC 11  
 |||||  
 Db 1 CKFLKCC 7

## RESULT 12

RI3887  
 ID RI3887 standard; Protein; 2749 AA.  
 AC RI3887;  
 DT 27-NOV-1991 (first entry)  
 DE Inositol-3-phosphate binding peptide.  
 KW I-3-P; transformation; inositol.  
 KN J03183482-A.  
 PN 09-AUG-1991.  
 PD 14-DEC-1989; 324256.  
 PF 14-DEC-1989; JP-324256.  
 PR 14-DEC-1989; JP-324256.  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 DR WPI; 91-277584/38.  
 DR N-PSDB; Q13593.  
 DT New polypeptide having binding affinity to inositol-3-phosphate -  
 PT prep. by culturing cell contg. recombinant plasmid comprising  
 PT DNA and vector DNA  
 PS Disclosure: Fig 2(1-3); l1pp; Japanese.  
 CC The sequence encoding this peptide may be included in a  
 CC plasmid/vector for transformation of a host cell and mass-prodn.  
 CC of the peptide.  
 SQ Sequence 2749 AA;

Query Match 61.3%; Score 38; DB 1; Length 2749;  
 Best Local Similarity 54.5%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11  
 :|:||||  
 Db 1316 VKAEGKFIKKC 1326

## RESULT 13

ID W00168 standard; Protein; 2713 AA.  
 AC W00168;  
 DT 16-NOV-1996 (first entry)  
 DE Human T cell inositol 1,4,5-trisphosphate receptor.  
 KW Inositol 1,4,5-trisphosphate receptor; calcium release channel;  
 KW IP3R; immunostimulant; immunosuppressive; T cell; T lymphocyte.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT region 318..332  
 FT /label= SI  
 FT /note= "alternatively spliced exon SI,  
 FT preferentially expressed in thymus  
 FT and spleen"  
 FT modified\_site 478..482  
 FT /label= Tyrosine\_phosphorylation\_site  
 FT modified\_site 1590  
 FT /label= PKA\_phosphorylation\_site  
 FT binding\_site 1689..1694  
 FT /label= Nucleotide\_binding\_site  
 FT modified\_site 1717  
 FT /label= PKA\_phosphorylation\_site

FT binding\_site 1737..1742  
 FT /label= Nucleotide\_binding\_site  
 FT region 1930..1939  
 FT /note= "region of high homology to rat type 1  
 FT IP3R"  
 FT binding\_site 1978..1983  
 FT /label= Nucleotide\_binding\_site  
 FT region 2000..2009  
 FT /note= "region of high homology to rat type 1  
 FT IP3R"  
 FT region 2240..2457  
 FT /label= M1  
 FT /note= "transmembrane region 1"  
 FT region 2270..2289  
 FT /label= M2  
 FT /note= "transmembrane region 2"  
 FT region 2314..2334  
 FT /label= M3  
 FT /note= "transmembrane region 3"  
 FT region 2357..2371  
 FT /label= M4  
 FT /note= "transmembrane region 4"  
 FT region 2404..2427  
 FT /label= M5  
 FT /note= "transmembrane region 5"  
 FT region 2533..2552  
 FT /label= M6  
 FT /note= "transmembrane region 6"  
 FT modified\_site 2615..2619  
 FT /label= Tyrosine\_phosphorylation\_site  
 FT peptide 2654..2665  
 FT /note= "immunogen (claim 13, page 45)"  
 PN W09624846-A1.  
 PD 15-AUG-1996.  
 PF 08-FEB-1996; U01735.  
 PR 09-FEB-1995; US-386039.  
 PA (MOUN ) MOUNT SINAI SCHOOL MEDICINE.  
 PI Marks AR;  
 DR WPI; 96-393018/39.  
 DR N-PSDB; T37090.  
 DT Human 1,4,5-tri-phosphate receptor/calcium release channel proteins  
 FT - used to produce assay systems for the identification and design of  
 FT agents which stimulate or suppress the immune system  
 PS Claim 1; Fig 1A-B; 57pp; English.  
 CC The amino acid sequence (W00168) of the type 1 inositol 1,4,5-  
 CC triphosphate receptor/calcium release channel (IP3R) was deduced  
 CC from a cDNA clone (T37090) isolated from a human leukaemic T-cell  
 CC cDNA library. IP3R tyrosine phosphorylation is associated with  
 CC T-cell activation. Such phosphorylation occurs in the vicinity of  
 CC a putative calcium channel region, and may be associated with the  
 CC opening of that channel. An alternatively spliced, neural-specific  
 CC SII exon is excluded from the present sequence. Type 1 IP3R can  
 CC be prep. by recombinant methods using prokaryotic or eukaryotic  
 CC host cells. It is used in assay systems to identify potential  
 CC immunostimulatory or immunosuppressive agents, and to raise  
 CC antibodies (see also W00169).  
 SQ Sequence 2713 AA;

Query Match 61.3%; Score 38; DB 1; Length 2713;  
 Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11  
 :|:||||  
 Db 1318 VKAEGKFIKKC 1328

## RESULT 14

W07632  
 ID W07632 standard; protein; 2695 AA.  
 AC W07632;  
 DT 04-FEB-1997 (first entry)

DE Human type I inositol triphosphate receptor.  
 KW Human; type I; inositol triphosphate; receptor; screen; antagonist;  
 OS Inhibition; calcium release.  
 PN Homo sapiens.  
 PD J08245698-A.  
 PF 24-SEP-1996.  
 PR 14-MAR-1995; 083142.  
 PR 14-MAR-1995; JP-083142.  
 PA (SOSE-) SOSEI KK.  
 DR WPI; 96-482258/48.  
 PT Human type I inositol triphosphate receptor - useful as a calcium  
 release inhibitor.  
 PS Claim 1; Pages 7-14; 16pp; Japanese.  
 CC The present sequence is that of the human type I inositol  
 triphosphate (IP3) receptor (IP3R1), which can be used to screen  
 for an IP3 like substance or an IP3 antagonist, or to inhibit  
 calcium release.  
 SQ Sequence 2695 AA;

Query Match 61.3%; Score 38; DB 1; Length 2695;  
 Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11  
 :| : ||:|  
 Db 1302 VKAGKFIKCC 1312

RESULT 15  
 W38732  
 ID W38732 standard; Protein; 37 AA.  
 AC W38732;  
 DT 10-NOV-1998 (first entry)  
 DE Streptococcus pneumoniae protein of unknown function.  
 KW Streptococcus pneumoniae protein; genetic immunisation; antagonist;  
 KW immunological response; inoculation; antibody production; inhibitor;  
 KW T cell immune response; antimicrobial compound; bacterial adhesion;  
 KW extracellular matrix protein; protein-mediated cell invasion; wound;  
 KW pathogenesis.  
 OS Streptococcus pneumoniae.  
 PN W09743303-A1.  
 PD 20-NOV-1997.  
 PF 14-MAY-1997; U07950.  
 PR 14-MAY-1996; US-017670.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI (SMIK ) SMITHKLINE BEECHAM PLC.  
 PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO,  
 PI Stodola RK;  
 DR WPI: 98-008793/01.  
 DR N-PSDB: T98774.  
 PT Novel Streptococcus pneumoniae proteins and related DNA - useful for  
 diagnosing anti-microbial agents for treatment of bacterial  
 infections  
 PS Claim 12; Pages 460-461; 483pp; English.  
 CC This sequence represents a Streptococcus pneumoniae protein of  
 unknown function, and is encoded by a DNA sequence of the invention.  
 CC The DNA sequences were isolated from Streptococcus pneumoniae strain  
 CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the  
 CC invention can be used to identify compounds which interact with and  
 CC inhibit or activate the activity of the proteins. Antagonists can be  
 CC used to treat diseases caused by S. pneumoniae proteins, through genetic  
 CC immunisation. They can also be used to induce an immunological response  
 CC in a mammal by inoculation with the S. pneumoniae proteins or delivery  
 CC of the encoding nucleic acids in a vector adequate to produce antibody  
 CC and/or T cell immune responses to protect the animal from disease. The  
 CC proteins can also be used to identify antimicrobial compounds which are  
 CC capable of inhibiting their bioactivity. In particular the proteins of  
 CC the invention can be used to prevent adhesion of bacteria to mammalian  
 CC extracellular matrix proteins on in-dwelling devices or in wounds, to  
 CC block protein-mediated mammalian cell invasion, and to block the normal  
 CC progression of pathogenesis in infections initiated other than by the  
 CC implantation of in-dwelling devices or other surgical techniques.

SQ Sequence 37 AA;  
 Query Match 61.3%; Score 38; DB 1; Length 37;  
 Best Local Similarity 70.0%; Pred. No. 5.6;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 KTKCKFLKCC 11  
 :| : ||:|  
 Db 20 RNKDKFLKCC 29

Search completed: September 7, 1999, 22:50:06  
 Job time: 7887 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:50 ; Search time 82.37 Seconds  
(without alignments)  
1.318 Million cell updates/sec

Title: US-09-124-280A-35  
Perfect score: 62  
Sequence: 1 IKTKCKFLKCC 11

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	11	1	US-08-049-871-6
2	62	100.0	11	1	US-07-819-893-6
3	62	100.0	11	1	US-08-280-397-6
4	62	100.0	11	2	US-08-456-112B-35
5	58	93.5	10	1	US-08-049-871-2
6	58	93.5	10	1	US-07-819-893-2
7	58	93.5	10	1	US-08-280-397-2
8	58	93.5	10	2	US-08-218-026-49
9	58	93.5	11	2	US-08-218-026-50
10	58	93.5	10	2	US-08-653-632-49
11	58	93.5	11	2	US-08-653-632-50
12	58	93.5	10	2	US-08-456-112B-31
13	43	69.4	7	1	US-08-049-871-1
14	43	69.4	7	1	US-07-819-893-1
15	43	69.4	7	1	US-08-280-397-1
16	43	69.4	7	2	US-08-456-112B-30
17	38	61.3	2713	3	PCT-US96-01735-1
18	37	59.7	10	1	US-08-097-830E-15
19	37	59.7	10	1	US-08-097-830E-17
20	37	59.7	10	2	US-08-456-112B-15
21	37	59.7	10	2	US-08-456-112B-17
22	34	54.8	553	1	US-08-328-322-10
23	34	54.8	509	1	US-08-328-322-19
24	34	54.8	1824	2	US-08-680-327-3
25	33.5	54.0	10	1	US-08-049-871-7
26	33.5	54.0	10	1	US-07-819-893-7
27	33.5	54.0	10	1	US-08-280-397-7
28	33.5	54.0	10	2	US-08-218-026-48
29	33.5	54.0	10	2	US-08-653-632-48
30	33.5	54.0	10	2	US-08-456-112B-36
31	33.5	54.0	10	3	PCT-US94-01234-46
32	33	53.2	1003	1	US-07-991-867B-6
33	33	53.2	10	1	US-08-097-830E-16
34	33	53.2	1003	1	US-08-107-755A-6
35	33	53.2	223	1	US-08-430-633-1
36	33	53.2	211	1	US-08-575-964-1
37	33	53.2	10	2	US-08-456-112B-16
38	33	53.2	391	2	US-08-644-034A-1
39	32	51.6	551	1	US-08-120-960-2

ALIGNMENTS

RESULT 1  
US-08-049-871-6  
; Sequence 6, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-6  
  
Query Match 100.0%; Score 62; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 IKTKCKFLKCC 11  
| | | | | | | | | | | |  
Db 1 IKTKCKFLKCC 11  
  
RESULT 2  
US-07-819-893-6  
; Sequence 6, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: circular  
US-07-819-893-6

Query Match 100.0%; Score 62; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKCKFLKKC 11  
| | | | | | | | | |  
Db 1 IKTKCKFLKKC 11

RESULT 3  
US-08-280-397-6  
;; Sequence 6, Application US/08280397  
;; Patent No. 5589459  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
US-08-280-397-6

Query Match 100.0%; Score 62; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKCKFLKKC 11  
| | | | | | | | | |  
Db 1 IKTKCKFLKKC 11

RESULT 4  
US-08-456-112B-35  
;; Sequence 35, Application US/08456112B  
;; Patent No. 5834430  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 35:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
US-08-456-112B-35

Query Match 100.0%; Score 62; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTCKFLKCC 11  
| | | | | | | | | |  
Db 1 IKTCKFLKCC 11

RESULT 5  
US-08-049-871-2  
; Sequence 2, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-2

Query Match 93.5%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
| | | | | | | | | |  
Db 1 KTKCKFLKCC 10

RESULT 6  
US-07-819-893-2  
; Sequence 2, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas

; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
US-07-819-893-2

Query Match 93.5%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
| | | | | | | | | |  
Db 1 KTKCKFLKCC 10

RESULT 7  
US-08-280-397-2  
; Sequence 2, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-2

Query Match 93.5%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
|||||  
Db 1 KTKCKFLKCC 10

RESULT 8  
US-08-218-026-49  
Sequence 49, Application US/08218026  
Patent No. 5786324  
GENERAL INFORMATION:  
APPLICANT: Gray, Beulah  
APPLICANT: Haseman, Judith R.  
TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
Activity and Endotoxin Neutralizing Activity for Gram  
TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5786324west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/218,026  
FILING DATE: 24-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalchuk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 600.286US01  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-218-026-49

Query Match 93.5%; Score 58; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
|||||  
Db 1 KTKCKFLKCC 10

RESULT 9  
US-08-218-026-50  
Sequence 50, Application US/08218026  
Patent No. 5786324  
GENERAL INFORMATION:  
APPLICANT: Gray, Beulah  
APPLICANT: Haseman, Judith R.  
APPLICANT: Mayo, Kevin  
TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
Activity and Endotoxin Neutralizing Activity for Gram  
TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5786324west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/218,026  
FILING DATE: 24-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalchuk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 600.286US01  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
MOLECULE TYPE: peptide  
US-08-218-026-50

Query Match 93.5%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
|||||  
Db 2 KTKCKFLKCC 11

RESULT 10  
US-08-653-632-49  
Sequence 49, Application US/08653632  
Patent No. 5830860  
GENERAL INFORMATION:  
APPLICANT: GRAY, Beulah  
APPLICANT: HASEMAN, Judith R.  
APPLICANT: MAYO, Kevin  
TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 5830860west Center, 90 South Seventh St  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
US-08-653-632-49



; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/653,632  
; FILING DATE: 24-MAY-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/218026  
; FILING DATE: 24-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kettelberger, Denise M  
; REGISTRATION NUMBER: 33,924  
; REFERENCE/DOCKET NUMBER: 600.286US11  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/371-5268  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; US-08-653-632-49

Query Match 93.5%; Score 58; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKKC 11  
| | | | |  
Db 1 KTKCKFLKKC 10

RESULT 11  
US-08-653-632-50  
; Sequence 50, Application US/08653632  
; Patent No. 5830860  
; GENERAL INFORMATION:  
; APPLICANT: GRAY, Beulah  
; APPLICANT: HASEMAN, Judith R.  
; APPLICANT: MAYO, Kevin  
; TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
; STREET: 3100 No. 5830860west Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/653,632  
; FILING DATE: 24-MAY-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/218026  
; FILING DATE: 24-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kettelberger, Denise M  
; REGISTRATION NUMBER: 33,924

; REFERENCE/DOCKET NUMBER: 600.286US11  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/371-5268  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; US-08-653-632-50

Query Match 93.5%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKKC 11  
| | | | |  
Db 2 KTKCKFLKKC 11

RESULT 12  
US-08-456-112B-31  
; Sequence 31, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: PORIO, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-31

Query Match 93.5%; Score 58; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKCKFLKCC 11  
 Db 1 KTKCKFLKCC 10

## RESULT 13

US-08-049-871-1  
 ; Sequence 1, Application US/08049871  
 ; Patent No. 5358933

## GENERAL INFORMATION:

APPLICANT: Porrio, Massimo  
 TITLE OF INVENTION: Synthetic Peptides for Detoxification  
 TITLE OF INVENTION: of Bacterial Endotoxins and for the  
 TITLE OF INVENTION: Prevention and Treatment of Septic  
 TITLE OF INVENTION: Shock  
 NUMBER OF SEQUENCES: 8

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

## COMPUTER:

IBM PS/2

## OPERATING SYSTEM:

Word Perfect 5.1

## SOFTWARE:

Word Perfect 5.1

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/049,871

## FILING DATE:

CLASSIFICATION: 514

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/658,744

## FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

TOPOLOGY: circular

US-08-049-871-1

Query Match 69.4%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CKFLKCC 11  
 Db 1 CKFLKCC 7

## RESULT 14

US-07-819-893-1  
 ; Sequence 1, Application US/07819893  
 ; Patent No. 5371186

## GENERAL INFORMATION:

APPLICANT: Porrio, Massimo  
 TITLE OF INVENTION: Synthetic Peptides for Detoxification  
 TITLE OF INVENTION: of Bacterial Endotoxins and for the  
 TITLE OF INVENTION: Prevention and Treatment of Septic  
 TITLE OF INVENTION: Shock  
 NUMBER OF SEQUENCES: 10

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036

STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

## COMPUTER:

IBM PS/2

## OPERATING SYSTEM:

Word Perfect 5.1

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/819,893

FILING DATE: 19920115

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: AMINO ACID

TOPOLOGY: circular

US-07-819-893-1

Query Match 69.4%; Score 43; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CKFLKCC 11  
 Db 1 CKFLKCC 7

## RESULT 15

US-08-280-397-1  
 ; Sequence 1, Application US/08280397  
 ; Patent No. 5589459

## GENERAL INFORMATION:

APPLICANT: Porrio, Massimo  
 TITLE OF INVENTION: Synthetic Peptides for Detoxification  
 TITLE OF INVENTION: of Bacterial Endotoxins and for the  
 TITLE OF INVENTION: Prevention and Treatment of Septic  
 TITLE OF INVENTION: Shock  
 NUMBER OF SEQUENCES: 10

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson & Costigan, P.C.

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

## COMPUTER:

IBM PS/2

## OPERATING SYSTEM:

Word Perfect 5.1

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/280,397

FILING DATE: 07/26/94

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/819,893

FILING DATE: 01/16/92

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-280-397-1

Query Match 69.4%; Score 43; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CKFLKNC 11  
Db 1 CKFLKNC 7

Search completed: September 7, 1999, 23:07:50  
Job time: 1748 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:49 ; Search time 105.14 Seconds  
(without alignments)  
4.192 Million cell updates/sec

Title: US-09-124-280A-35  
Perfect score: 62  
Sequence: 1 IKTKCKFLKCC 11

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	62.9	73	2 C44088	homeotic protein H
2	38.5	62.1	239	2 T03078	conserved hypothet
3	38	61.3	2749	1 ACMSIT	inositol 1,4,5-tri
4	38	61.3	2734	2 B36579	inositol 1,4,5-tri
5	38	61.3	2749	2 A36579	inositol 1,4,5-tri
6	38	61.3	2713	2 A55713	inositol 1,4,5-tri
7	38	61.3	2695	2 S54974	type 1 inositol 1,
8	38	61.3	347	2 A70173	fema protein (fema
9	38	61.3	206	2 S49882	hypothetical prote
10	37	59.7	1121	2 S54504	hypothetical prote
11	37	59.7	554	2 JU0188	membrane protein 4
12	36	58.1	491	2 B71957	ATP-dependent RNA
13	36	58.1	492	2 G64550	ATP-dependent RNA
14	36	58.1	445	2 G01452	NIMA-like protein
15	36	58.1	1170	2 A56157	chromosome segrega
16	36	58.1	244	2 A49202	cysteine-rich secr
17	36	58.1	1766	2 S03701	141K protein - pea
18	36	58.1	170	2 C71199	hypothetical prote
19	35	56.5	2670	2 A46719	inositol 1,4,5-tri
20	35	56.5	2671	2 A49873	inositol 1,4,5-tri
21	35	56.5	1059	2 B24872	retrovirus-related
22	35	56.5	1040	2 S34639	pol protein - frui
23	35	56.5	449	2 C71959	glutamyl tRNA redu
24	35	56.5	60	2 JS0583	hypothetical 7K pr
25	35	56.5	135	2 C70028	conserved hypothet
26	35	56.5	1122	2 F71606	probable multiple
27	34	54.8	1089	2 S33727	platelet-derived g
28	34	54.8	335	2 A25968	DNA-directed RNA p
29	34	54.8	553	2 S16809	phosphoprotein pho
30	34	54.8	298	2 D69023	cell division inh
31	34	54.8	246	2 A64326	hypothetical prote
32	34	54.8	859	2 H70327	DNA mismatch repai
33	34	54.8	644	2 A42220	helix-loop-helix p
34	34	54.8	921	2 S49965	probable membrane
35	34	54.8	984	2 S14382	hypothetical prote
36	34	54.8	403	2 S07825	hypothetical prote
37	34	54.8	603	2 T00379	KIAA0640 protein -
38	34	54.8	189	2 G70485	hypothetical prote
39	34	54.8	128	2 F71170	hypothetical prote

40 33.5 54.0 936 2 S64384  
41 33 53.2 356 1 DEFUGW  
42 33 53.2 223 1 QQBE26  
43 33 53.2 1003 1 PYVZAM  
44 33 53.2 505 2 JS0722  
45 33 53.2 282 2 B71444

probable membrane  
malate dehydrogena  
B2LF2 protein - hu  
spheroidin precurs  
cytochrome P450 AL  
hypothetical prote

ALIGNMENTS

RESULT 1

C44088  
homeotic protein HAR22 - Arabidopsis thaliana (fragments)  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 02-Jul-1998  
C:Accession: C44088  
R:Schena, M.; Davis, R.W.  
Proc. Natl. Acad. Sci. U.S.A. 89, 3894-3898, 1992  
A:Title: HD-2ip proteins: members of an Arabidopsis homeodomain protein superfamily.  
A:Reference number: A44088; MUID:92237275  
A:Accession: C44088  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-73 <SCH>  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
F:18-32/Domain: homeobox homology (fragment) <HOX>

Query Match 62.9%; Score 39; DB 2; Length 73;  
Best Local Similarity 85.7%; Pred. No. 4.6;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 CKFLKCC 11  
|:|||||  
DB 39 CEFLKCC 45

RESULT 2

T03078  
conserved hypothetical protein 061L - Chilo iridescent virus  
C:Species: Chilo iridescent virus  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C:Accession: T03078  
R:Bahr, U.; Tidona, C.A.; Darai, G.  
Virus Genes 15, 235-245, 1997  
A:Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.  
A:Reference number: Z14834  
A:Accession: T03078  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-239 <BAH>  
A:Cross-references: EMBL:AF003534; NID:g2738385; PID:g2738425

Query Match 62.1%; Score 38.5; DB 2; Length 239;  
Best Local Similarity 63.6%; Pred. No. 14;  
Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 2 KTR-CKFLKCC 11  
||:|||||  
DB 83 KTRCKYVKKC 93

RESULT 3

ACMSIT  
inositol 1,4,5-trisphosphate receptor - mouse  
N:Alternate names: inositol-1,4,5-trisphosphate-binding protein P400; membrane-associa  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 02-Jul-1998  
C:Accession: S04844; S06796; C34955  
R:Furuichi, T.; Yoshikawa, S.; Mikoshiba, K.

Nucleic Acids Res. 17, 5385-5386, 1989  
 A:Title: Nucleotide sequence of cDNA encoding P400 protein in the mouse cerebellum.  
 A:Reference number: S04844; MUID:89345101  
 A:Accession: S04844

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-2749 <FUR>

A:Cross-references: EMBL:X15373; NID:G53568; PID:G53569

A:Experimental source: strain ICR; cerebellum

R:Furuichi, T.; Yoshikawa, S.; Miyawaki, A.; Wada, K.; Maeda, N.; Mikoshiba, K.

Nature 342, 32-38, 1989

A:Title: Primary structure and functional expression of the inositol 1,4,5-trisphosphate

A:Reference number: S06796; MUID:90044039

A:Accession: S06796

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-2749 <FU2>

A:Cross-references: EMBL:X15373; NID:G53568; PID:G53569

A:Experimental source: strain ICR; cerebellum

R:Nordquist, D.F.; Kozak, C.A.; Orr, H.T.

J. Neurosci. 8, 4780-4789, 1988

A:Title: cDNA cloning and characterization of three genes uniquely expressed in cerebellum

A:Reference number: A92975; MUID:89068131

A:Accession: C34955

A:Molecule type: mRNA

A:Residues: 2250-2674, 'P', 2676-2749 <NOR>

C:Superfamily: inositol-trisphosphate receptor

C:Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein

F:890-907/Domain: transmembrane #status predicted <TM1>

F:1960-1976/Domain: transmembrane #status predicted <TM2>

F:2276-2294/Domain: transmembrane #status predicted <TM3>

F:2308-2326/Domain: transmembrane #status predicted <TM4>

F:2334-2350/Domain: transmembrane #status predicted <TM5>

F:2352-2372/Domain: transmembrane #status predicted <TM6>

F:2391-2407/Domain: transmembrane #status predicted <TM7>

F:2440-2462/Domain: transmembrane #status predicted <TM8>

F:2570-2589/Domain: transmembrane #status predicted <TM9>

F:1598,1755/Binding site: phosphate (ser) (covalent) #status predicted

F:2475,4503,2622,2710/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match 61.3%; Score 38; DB 1; Length 2749;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IKTKCKFLKCC 11

:|: ||: |||

Db 1316 VKAEGKFIKKC 1326

RESULT 4

B36579

inositol 1,4,5-trisphosphate receptor 2 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 08-Sep-1997

C:Accession: B36579

R:Mignery, G.A.; Newton, C.L.; Archer III, B.T.; Suedhof, T.C.

J. Biol. Chem. 265, 12679-12685, 1990

A:Title: Structure and expression of the rat inositol 1,4,5-trisphosphate receptor.

A:Reference number: A36579; MUID:90324264

A:Accession: B36579

A:Molecule type: mRNA

A:Residues: 1-2734 <MIG>

A:Cross-references: GB:J05510; NID:G204673; PID:G204674

C:Superfamily: inositol-trisphosphate receptor

Query Match 61.3%; Score 38; DB 2; Length 2734;

Best Local Similarity 54.5%; Pred. No. 1.2e+02;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IKTKCKFLKCC 11

Db 1301 VKAEGKFIKKC 1311

:|: ||: |||

RESULT 5

A36579

inositol 1,4,5-trisphosphate receptor 1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 08-Sep-1997

C:Accession: A36579

R:Mignery, G.A.; Newton, C.L.; Archer III, B.T.; Suedhof, T.C.

J. Biol. Chem. 265, 12679-12685, 1990

A:Title: Structure and expression of the rat inositol 1,4,5-trisphosphate receptor.

A:Reference number: A36579; MUID:90324264

A:Accession: A36579

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2749 <MIG>

A:Cross-references: GB:J05510; NID:G204673; PID:G204675

C:Superfamily: inositol-trisphosphate receptor

C:Keywords: phosphoprotein; transmembrane protein

Query Match 61.3%; Score 38; DB 2; Length 2749;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IKTKCKFLKCC 11

:|: ||: |||

Db 1316 VKAEGKFIKKC 1326

RESULT 6

A55713

inositol 1,4,5-trisphosphate receptor type 1 - human

C:Species: Homo sapiens (man)

C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 24-Sep-1998

C:Accession: A55713

R:Harnick, D.J.; Jayaraman, T.; Ma, Y.; Mulieri, P.; Go, L.O.; Marks, A.R.

J. Biol. Chem. 270, 2833-2840, 1995

A:Title: The human type 1 inositol 1,4,5-trisphosphate receptor from T lymphocytes. S

A:Reference number: A55713; MUID:95155356

A:Accession: A55713

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2713 <HAR>

A:Cross-references: GB:L38019; NID:G1464750; PID:G598181

A:Note: authors translated the codon AAG for residue 2325 as Asn

C:Genetics:

A:Gene: GDB:ITPRL

A:Cross-references: GDB:127552; OMIM:147265

A:Map position: 3pter-3p26

C:Superfamily: inositol-trisphosphate receptor

Query Match 61.3%; Score 38; DB 2; Length 2713;

Best Local Similarity 54.5%; Pred. No. 1.2e+02;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IKTKCKFLKCC 11

:|: ||: |||

Db 1318 VKAEGKFIKKC 1328

RESULT 7

S54974

Type 1 inositol 1,4,5-trisphosphate receptor - human

C:Species: Homo sapiens (man)

C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 08-Sep-1997

C:Accession: S54974

R:Yamada, N.; Makino, Y.; Clark, R.A.; Pearson, D.W.; Mattei, M.G.; Guenet, J.L.; Oha

Blochem. J. 302, 781-790, 1994

A:Title: Human inositol 1,4,5-trisphosphate type-1 receptor, InSP(3)R1: structure, fu

A:Reference number: S54974  
A:Accession: S54974  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2695 <YAM>  
A:Cross-references: EMBL:D26070; NID:g559322; PID:d1005606; PID:g559323  
C:Superfamily: inositol-trisphosphate receptor

Query Match 61.3%; Score 38; DB 2; Length 2695;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTCKFLKCK 11  
: | : ||| | | |  
DB 1302 VRAEGKFLKCK 1312

RESULT 8  
A70173  
fema protein (fema) homolog - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
C:Accession: A70173  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
son, D.; Peterson, J.; Keriavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: A70173  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-347 <KLE>  
A:Cross-references: GB:AE001160; GB:AE000783; NID:g2688505; PID:g2688511; TIGR:BB0586  
A:Experimental source: strain B31

Query Match 61.3%; Score 38; DB 2; Length 347;  
Best Local Similarity 70.0%; Pred. No. 23;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTCKFLKCK 10  
: ||| | | | |  
DB 120 LTKIKYLKCK 129

RESULT 9  
S4982  
hypothetical protein YII127c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YI8277.02c  
C:Species: Saccharomyces cerevisiae  
C:Date: 13-Jan-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Nov-1997  
C:Accession: S4982  
R:Hamlyn, N.; Churcher, C.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: S4981  
A:Accession: S4982  
A:Molecule type: DNA  
A:Residues: 1-206 <HAM>  
A:Cross-references: GB:247047; EMBL:246833; NID:g603997; PID:g763219; MIPS:YII127c  
C:Genetics:  
A:Map position: 9L

Query Match 61.3%; Score 38; DB 2; Length 206;  
Best Local Similarity 70.0%; Pred. No. 15;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTCKFLKCK 10  
: | | | | | | |  
DB 67 IKKKCKLVKK 76

RESULT 10  
S54504  
hypothetical protein YPR030w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YP9367.10  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 21-Nov-1997  
C:Accession: S54504  
R:Badcock, K.; Churcher, C.M.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54059  
A:Accession: S54504  
A:Molecule type: DNA  
A:Residues: 1-1121 <BAD>  
A:Cross-references: EMBL:249274; NID:g809585; PID:g809595; MIPS:YPR030w  
A:Experimental source: strain AB972  
C:Genetics:  
A:Map position: 16R

Query Match 59.7%; Score 37; DB 2; Length 1121;  
Best Local Similarity 63.6%; Pred. No. 86;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IKTCKFLKCK 11  
: ||| | | | |  
DB 773 VKTKLKFPKYC 783

RESULT 11  
JU0188  
membrane protein 4.1 homolog - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Feb-1994 #sequence\_revision 20-Aug-1994 #text\_change 21-Nov-1998  
C:Accession: I55505; JU0188  
R:Takeuchi, K.; Kawashima, A.; Nagafuchi, A.; Tsukita, S.  
J. Cell Sci. 107, 1921-1928, 1994  
A:Title: Structural diversity of band4.1 superfamily members.  
A:Reference number: I55505; MUID:95074267  
A:Accession: I55505  
A:Status: preliminary; translated from GB/EMBL/DBBJ  
A:Molecule type: mRNA  
A:Residues: 1-554 <RES>  
A:Cross-references: GB:D28818; NID:g466547; PID:g466548  
C:Superfamily: protein 4.1 membrane-binding domain homology  
F.13-295/Domain: protein 4.1 membrane-binding domain homology <B41>

Query Match 59.7%; Score 37; DB 2; Length 554;  
Best Local Similarity 70.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KTKCKFLKCK 11  
: | | | | | | |  
DB 278 KTKCKHLWKC 287

RESULT 12  
B71957  
ATP-dependent RNA helicase dead - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 05-Mar-1999  
C:Accession: B71957  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.  
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MUID:99120557  
A:Accession: B71957  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-491 <ARN>  
 A:Cross-references: GB:AE001461; GB:AE001439; NID:g4154749; PID:g4154758  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: dead  
 C:Superfamily: unassigned DEAD/H box helicases; DEAD/H box helicase homology

Query Match 58.1%; Score 36; DB 2; Length 491;  
 Best Local Similarity 50.0%; Pred. No. 65;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKCKFLKK 10  
 :|:|:|:|  
 Db 127 VKQCEFIKK 136

RESULT 13  
 G64550  
 ATP-dependent RNA helicase, DEAD-box family - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 05-Jun-1998  
 C:Accession: G64550  
 R:Tomb, J.F.; White, O.; Kervilave, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: G64550  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-492 <TOM>  
 A:Cross-references: GB:AE000544; GB:AE000511; NID:g2313337; PID:g2313340; TIGR:HP0247  
 C:Superfamily: unassigned DEAD/H box helicases; DEAD/H box helicase homology  
 C:Keywords: ATP; P-loop  
 F:64-350/Domain: DEAD/H box helicase homology <DEAD>  
 F:64-71/Region: nucleotide-binding motif A (P-loop)  
 F:164-169/Region: nucleotide-binding motif B  
 F:168-171/Region: DEAD motif

Query Match 58.1%; Score 36; DB 2; Length 492;  
 Best Local Similarity 50.0%; Pred. No. 65;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKCKFLKK 10  
 :|:|:|:|  
 Db 127 VKQCEFIKK 136

RESULT 14  
 G01452  
 NIMA-like protein kinase 1 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 17-Jul-1998  
 C:Accession: G01452  
 R:Lu, K.P.  
 submitted to the EMBL Data Library, June 1994  
 A:Reference number: G07172  
 A:Accession: G01452  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-445 <LUJ>  
 A:Cross-references: EMBL:U11050; NID:g507874; PID:g507875  
 C:Genetics:  
 A:Gene: NIK1  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology  
 F:6-271/Domain: protein kinase homology <KIN>  
 F:14-22/Region: protein kinase ATP-binding motif

Query Match 58.1%; Score 36; DB 2; Length 445;  
 Best Local Similarity 77.8%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKCKFLKK 10  
 :|:|:|:|  
 Db 407 KSKCKDLKK 415

RESULT 15  
 A56157  
 chromosome segregation protein SMC2 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: DA-box protein; protein R003; protein YFR031c; structural maintena  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 17-Mar-1999  
 C:Accession: A56157; S56286; S62242; S63836; S48530  
 R:Strunnikov, A.V.; Hogan, E.; Koshland, D.  
 Genes Dev. 9, 587-599, 1995  
 A:Title: SMC2, a Saccharomyces cerevisiae gene essential for chromosome segregation a  
 A:Reference number: A56157; MUID:95212908  
 A:Accession: A56157  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1170 <STR>  
 A:Cross-references: GB:U05820; NID:g468039; PID:g468040  
 R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas  
 submitted to the EMBL Data Library, May 1995  
 A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces  
 A:Reference number: S56186  
 A:Accession: S56286  
 A:Molecule type: DNA  
 A:Residues: 1-1170 <MUR>  
 A:Cross-references: EMBL:D50617; NID:g836685; PID:d1009911; PID:g836786; MIPS:YFR031c  
 R:Murakami, Y.  
 submitted to the EMBL Data Library, December 1994  
 A:Reference number: S62230  
 A:Accession: S62242  
 A:Molecule type: DNA  
 A:Residues: 1-1170 <MUW>  
 A:Cross-references: EMBL:D44602; NID:g893419; PID:d1008633; PID:g893426  
 R:Eki, T.; Naitou, M.; Hagiwara, H.; Abe, M.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.  
 Yeast 12, 177-190, 1996  
 A:Title: Fifteen open reading frames in a 30.8 kb region of the right arm of chromoso  
 A:Reference number: S63830; MUID:96287654  
 A:Accession: S63836  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1170 <EKI>  
 A:Cross-references: EMBL:D44602; NID:g893419; PID:d1008633; PID:g893426  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995  
 C:Genetics:  
 A:Gene: SGD:SMC2  
 A:Cross-references: SGD:S0001927; MIPS:YFR031c  
 A:Map position: 6R  
 C:Superfamily: conserved hypothetical P115 protein  
 C:Keywords: transmembrane protein  
 F:41-57/Domain: transmembrane #status predicted <TM1>  
 F:1090-1106/Domain: transmembrane #status predicted <TM2>

Query Match 58.1%; Score 36; DB 2; Length 1170;  
 Best Local Similarity 72.7%; Pred. No. 1.3e+02;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IKTKCKFLKK 11  
 :|:|:|:|  
 Db 761 IKTKQMSLKK 771

Search completed: September 7, 1999, 23:22:50  
 Job time: 957 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:15 ; Search time 71.09 Seconds  
(without alignments)  
4.374 Million cell updates/sec

Title: US-09-124-280A-35  
Perfect score: 62  
Sequence: 1 IKTKCKFLKCC 11

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	62.9	274	1	HAT9_ARATH
2	39	62.9	278	1	HT22_ARATH
3	38	61.3	2749	1	IP3R_MOUSE
4	38	61.3	2749	1	IP3R_RAT
5	38	61.3	206	1	YIN7_YEAST
6	37	59.7	554	1	NBL4_MOUSE
7	36	58.1	244	1	AEGL1_MOUSE
8	36	58.1	445	1	NEK2_HUMAN
9	36	58.1	1170	1	SMC2_YEAST
10	35	56.5	1059	1	POL2_DROME
11	34	54.8	632	1	FBF2_CAEEL
12	34	54.8	353	1	MDHG_SOYBN
13	34	54.8	859	1	MUTS_AQUAE
14	34	54.8	403	1	NOF2_DROME
15	34	54.8	984	1	NOF_DROME
16	34	54.8	553	1	P2B1_YEAST
17	34	54.8	1089	1	PGDS_MOUSE
18	34	54.8	335	1	RPC5_YEAST
19	34	54.8	246	1	Y208_METJA
20	34	54.8	921	1	YIB7_YEAST
21	33.5	54.0	936	1	YG2K_YEAST
22	33	53.2	220	1	ANT4_HYDMA
23	33	53.2	218	1	ATH4_ARATH
24	33	53.2	191	1	CC42_DROME
25	33	53.2	505	1	CP5T_CANNA
26	33	53.2	282	1	HAT1_ARATH
27	33	53.2	208	1	HAT2_ARATH
28	33	53.2	315	1	HAT3_ARATH
29	33	53.2	284	1	HAT4_ARATH
30	33	53.2	165	1	HT14_ARATH
31	33	53.2	88	1	MAT8_MOUSE
32	33	53.2	356	1	MDHG_CITVU
33	33	53.2	356	1	MDHG_CUCSA
34	33	53.2	3344	1	POLG_PRSVH
35	33	53.2	38	1	SKL1_ORTSC
36	33	53.2	1002	1	SPHR_AMEPV
37	33	53.2	525	1	SYH_CAEEL
38	33	53.2	798	1	YLFN_CAEEL
39	33	53.2	223	1	Y2L2_EBV
40	32.5	52.4	100	1	TAT_SIVAT
41	32	51.6	1700	1	BAR3_CHITE
42	32	51.6	24	1	BRI_RANBP
43	32	51.6	1863	1	BRC1_HUMAN

44 32 51.6 550 1 CBS\_HUMAN P35520 homo saplen  
45 32 51.6 560 1 CBS\_RAT P32232 rattus norv

ALIGNMENTS

```
RESULT 1
HAT9_ARATH
ID HAT9_ARATH STANDARD; PRT; 274 AA.
AC P46603;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HOMOBOX-LEUCINE ZIPPER PROTEIN HAT9 (HD-ZIP PROTEIN 9).
GN HAT9.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA MEDLINE; 94359937.
RA SCHENA M., DAVIS R.W.;
RT "Structure of homeobox-leucine zipper genes suggests a model for the
    evolution of gene families.";
RL PROC. NATL. ACAD. SCI. U.S.A. 91:8393-8397(1994).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
    the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
-----
CC EMBL: U09341; G549892; -
CC EMBL: U09342; G549893; -
CC PROSITE: PS00027; HOMOBOX_1; 1.
CC PROSITE: PS50071; HOMOBOX_2; 1.
CC PFAM: PF00046; homeobox; 1.
CC HSSP: P01366; IAKH.
CC HOMOBOX; DNA-BINDING: NUCLEAR PROTEIN.
FT DNA_BIND 110 169 HOMOBOX.
FT DOMAIN 177 198 LEUCINE-ZIPPER.
FT DOMAIN 90 93 POLY-GLU.
FT DOMAIN 226 240 POLY-GLY.
FT CONFLICT 90 90 E -> V (IN G549893).
SQ SEQUENCE 274 AA; 29867 MW; DA3C3355 CRC32;

Query Match 62.9%; Score 39; DB 1; Length 274;
Best Local Similarity 85.7%; Pred. No. 7.3;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 CKFLKCC 11
DB 174 CEFLKCC 180

RESULT 2
HT22_ARATH
ID HT22_ARATH STANDARD; PRT; 278 AA.
AC P46604;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HOMOBOX-LEUCINE ZIPPER PROTEIN HAT22 (HD-ZIP PROTEIN 22).
GN HAT22.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
```

OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLOPHITES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, COLUMBIA;  
 RX MEDLINE: 94359937;  
 RA SCHENA M., DAVIS R.W.;  
 RT "Structure of homeobox-leucine zipper genes suggests a model for the  
 evolution of gene families.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:8393-8397(1994).  
 RN [2]  
 RP SEQUENCE OF 166-221 FROM N.A.  
 RC STRAIN=CV, COLUMBIA;  
 RX MEDLINE: 92237275;  
 RA SCHENA M., DAVIS R.W.;  
 RT "HD-ZIP proteins: members of an Arabidopsis homeodomain protein  
 superfamily.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:3894-3898(1992).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CC -!- SIMILARITY: STRONG, WITH OTHER HD-ZIP PLANT PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U09336; G549887;  
 DR EMBL: U09337; G549888;  
 DR EMBL: M90417; G166756;  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 DR PROSITE: PS00071; HOMEBOX\_2; 1.  
 DR PFAM: PF00046; homeobox; 1.  
 DR HSSP: P01366; 1YRN.  
 KW HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN.  
 FT DNA\_BIND 123 182  
 FT DOMAIN 190 211  
 FT CONFLICT 221 221 Y -> S (IN REF. 2).  
 SQ SEQUENCE 278 AA; 30729 MW; D50B698D CRC32;  
 -----  
 Query Match 62.98; Score 39; DB 1; Length 278;  
 Best Local Similarity 85.7%; Pred. No. 7.4;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 CKFLKCK 11  
 I:|||||  
 Db 187 CEFLKCK 193  
 -----  
 RESULT 3  
 ID IP3R\_MOUSE STANDARD; PRT; 2749 AA.  
 AC P11881;  
 DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN TYPE 1 RECEPTOR  
 DE (INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN P(400)) (TYPE 1 INSP3  
 DE RECEPTOR).  
 GN INSP3R OR ITPR1.  
 OS MUS MUSCULUS (MUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PURKINJE CELLS;  
 RX MEDLINE: 90044039;  
 RA FURUICHI T., YOSHIKAWA S., MIYAWAKI A., WADA K., MAEDA N.,  
 RA MIKOSHIBA K.;

\*Primary structure and functional expression of the inositol 1,4,5-  
 trisphosphate-binding protein P400.";  
 NATURE 342:32-38(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ICR; TISSUE=CEREBELLUM;  
 RX MEDLINE: 89345101;  
 RA FURUICHI T., YOSHIKAWA S., MIKOSHIBA K.;  
 RT "Nucleotide sequence of cDNA encoding P400 protein in the mouse  
 cerebellum.";  
 RL NUCLEIC ACIDS RES. 17:5385-5386(1989).  
 RN [3]  
 RP SEQUENCE OF 318-332 AND 1692-1731 FROM N.A., AND ALTERNATIVE SPLICING.  
 RC STRAIN=ICR;  
 RX MEDLINE: 91269797;  
 RA NAKAGAWA T., OKANO H., FURUICHI T., ARUGA J., MIKOSHIBA K.;  
 RT "The subtypes of the mouse inositol 1,4,5-trisphosphate receptor are  
 expressed in a tissue-specific and developmentally specific manner.";  
 PROC. NATL. ACAD. SCI. U.S.A. 88:6244-6248(1991).  
 CC -!- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND  
 MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
 CC THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL  
 EXTREMITY. ITS LARGE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE  
 BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE  
 PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
 RETICULUM.  
 CC -!- PTM: PHOSPHORYLATED BY CYCLIC-AMP KINASE. PHOSPHORYLATION PREVENTS  
 THE LIGAND-INDUCED OPENING OF THE CALCIUM CHANNELS.  
 CC -!- CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE RECEPTOR, MOST  
 PROBABLY BY INTERACTING WITH A DISTINCT CALCIUM-BINDING PROTEIN  
 WHICH THEN INHIBITS THE RECEPTOR.  
 CC -!- ALTERNATIVE PRODUCTS: ADDITIONAL SUBTYPES OF INSP3R ARISE BY  
 ALTERNATIVE SPLICING OF THE SAME GENE.  
 CC -!- SIMILARITY: TO RYANODINE RECEPTOR.  
 CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X15373; G53569;  
 DR EMBL: M75986; G198412;  
 DR EMBL: M75987; G554171;  
 DR PIR: S04844; ACMSIT.  
 DR MGD: MGI:96623; ITPR1.  
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; PHOSPHORYLATION;  
 KW ENDOPLASMIC RETICULUM; IONIC CHANNEL; ION TRANSPORT; CALCIUM CHANNEL;  
 KW ALTERNATIVE SPLICING.  
 FT DOMAIN 1 2273  
 FT TRANSMEM 2274 2294  
 FT TRANSMEM 2308 2326  
 FT TRANSMEM 2334 2356  
 FT TRANSMEM 2365 2387  
 FT TRANSMEM 2391 2407  
 FT TRANSMEM 2440 2462  
 FT TRANSMEM 2530 2549  
 FT TRANSMEM 2570 2589  
 FT DOMAIN 2590 2749  
 FT VARSPIC 318 332  
 FT VARSPIC 1592 1731  
 FT VARSPIC 1715 1715  
 FT VARSPIC 1715 1731  
 FT MOD\_RES 1588 1588  
 FT MOD\_RES 1755 1755  
 SQ SEQUENCE 2749 AA; 313193 MW; B19B78F2 CRC32;

Query Match

61.3%; Score 38; DB 1; Length 2749;

Best Local Similarity 54.58; Pred. No. 73; Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKCKFLKKC 11  
:|:|:|

Db 1316 VKAEGKFKKC 1326

RESULT 4

IP3R\_RAT  
ID IP3R\_RAT STANDARD; PRT; 2749 AA.  
AC P29994;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN TYPE 1 RECEPTOR  
DE (INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN P(400)) (TYPE 1 INSP3 RECEPTOR).  
GN INSP3R.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90324284.  
RA MIGNERY G.A., NEWTON C.L., ARCHER B.T. III, SUEHOF T.C.;  
RT "Structure and expression of the rat inositol 1,4,5-trisphosphate receptor.";  
RL J. BIOL. CHEM. 265:12679-12685(1990).  
RN [2]  
RP SEQUENCE OF 1652-1824 FROM N.A., AND ALTERNATIVE SPLICING.  
RC TISSUE-BRAIN;  
RX MEDLINE; 91187909.  
RA DANOFF S.K., FERRIS C.D., DONATH C., FISCHER G.A., MUNEMITSU S.,  
RA ULLRICH A., SNYDER S.H., ROSS C.A.;  
RT "Inositol 1,4,5-trisphosphate receptors: distinct neuronal and nonneuronal forms derived by alternative splicing differ in phosphorylation.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 88:2951-2955(1991).  
RC [-]  
CC FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
CC THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
CC [-]  
CC SUBUNIT: HOMOTETRAMER.  
CC [-]  
CC SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC RETICULUM.  
CC [-]  
CC TISSUE SPECIFICITY: THE LONGER ISOFORM IS FOUND IN THE BRAIN WHILE THE SHORTER ISOFORM IS FOUND IN THE FETAL BRAIN AND PERIPHERAL TISSUES.  
CC [-]  
CC PTM: PHOSPHORYLATED BY CYCLIC-AMP KINASE. PHOSPHORYLATION PREVENTS THE LIGAND-INDUCED OPENING OF THE CALCIUM CHANNELS.  
CC [-]  
CC CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE RECEPTOR, MOST PROBABLY BY INTERACTING WITH A DISTINCT CALCIUM-BINDING PROTEIN WHICH THEN INHIBITS THE RECEPTOR.  
CC [-]  
CC ALTERNATIVE PRODUCTS: TWO FORMS, A LONG (NEURONAL) AND A SHORT (NONNEURONAL) FORM, ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.  
CC [-]  
CC SIMILARITY: TO RYANODINE RECEPTOR.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J05510; G204675; -;  
CC EMBL; J05510; G204674; -;  
CC EMBL; M64699; G204967; -;  
CC EMBL; M64698; G204969; -;

DR PIR; A36579; A36579.  
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; PHOSPHORYLATION;  
KW ENDOPLASMIC RETICULUM; IONIC CHANNEL; ION TRANSPORT; CALCIUM CHANNEL;  
KW ALTERNATIVE SPLICING.  
FT DOMAIN 1 2273 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2274 2294 M1 (POTENTIAL).  
FT TRANSMEM 2308 2326 M2 (POTENTIAL).  
FT TRANSMEM 2334 2356 M3 (POTENTIAL).  
FT TRANSMEM 2365 2387 M4 (POTENTIAL).  
FT TRANSMEM 2391 2407 M5 (POTENTIAL).  
FT TRANSMEM 2440 2462 M6 (POTENTIAL).  
FT TRANSMEM 2530 2549 M7 (POTENTIAL).  
FT TRANSMEM 2570 2589 M8 (POTENTIAL).  
FT DOMAIN 2590 2749 CYTOPLASMIC (POTENTIAL).  
FT MOD\_RES 1589 1589 PHOSPHORYLATION (BY CAPK) (POTENTIAL).  
FT MOD\_RES 1755 1755 PHOSPHORYLATION (BY CAPK) (POTENTIAL).  
FT VARIANT 1372 1372 MISSING (IN ALL, BUT P17 CLONES).  
FT VARSPLIC 322 336 MISSING (IN VARIANT).  
FT VARSPLIC 1693 1731 MISSING (IN SHORT FORM).  
FT CONFLICT 1716 1716 E -> EQ (IN REF. 2).  
FT CONFLICT 1763 1763 P -> R (IN REF. 2).  
SQ SEQUENCE 2749 AA; 313132 MW; A64DE4B1 CRC32;

Query Match 61.3%; Score 38; DB 1; Length 2749;  
Best Local Similarity 54.5%; Pred. No. 73;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKCKFLKKC 11  
:|:|:|

Db 1316 VKAEGKFKKC 1326

RESULT 5

YIM7\_YEAST  
ID YIM7\_YEAST STANDARD; PRT; 206 AA.  
AC P40470;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 23.8 KD PROTEIN IN MET18-STH1 INTERGENIC REGION. YIL127C.  
GN YIL127C.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA BARRELL B.G., BADCOCK K., BANKIER A.T., BOWMAN S., BROWN D.,  
RA CHURCHER C.M., CONNOR R., COPSEY T., DEAR S., DEVLIN K., FRASER A.,  
RA GENTLES S., HAMLYN N., HORSNELL T.S., HUNT S., JAGELS K., JONES M.,  
RA LOUIS E., LYE G., MOULE S., MOULE T., ODELL C., PEARSON D.,  
RA RAJANDREAM M.A., RILES L., ROWLEY N., SKELTON J., SMITH V.,  
RA WALSH S.V., WHITEHEAD S.;  
RC SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
RL  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z47047; G763219; -;  
CC EMBL; Z46833; G599990; -;  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 206 AA; 23845 MW; A140E520 CRC32;

Query Match 61.3%; Score 38; DB 1; Length 206;  
Best Local Similarity 70.0%; Pred. No. 8.4;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Best Local Similarity 54.58; Pred. No. 73; Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKCKFLKKC 11  
:|:|:|

Db 1316 VKAEGKFKKC 1326

RESULT 4

IP3R\_RAT  
ID IP3R\_RAT STANDARD; PRT; 2749 AA.  
AC P29994;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN TYPE 1 RECEPTOR  
DE (INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN P(400)) (TYPE 1 INSP3 RECEPTOR).  
GN INSP3R.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90324284.  
RA MIGNERY G.A., NEWTON C.L., ARCHER B.T. III, SUEHOF T.C.;  
RT "Structure and expression of the rat inositol 1,4,5-trisphosphate receptor.";  
RL J. BIOL. CHEM. 265:12679-12685(1990).  
RN [2]  
RP SEQUENCE OF 1652-1824 FROM N.A., AND ALTERNATIVE SPLICING.  
RC TISSUE-BRAIN;  
RX MEDLINE; 91187909.  
RA DANOFF S.K., FERRIS C.D., DONATH C., FISCHER G.A., MUNEMITSU S.,  
RA ULLRICH A., SNYDER S.H., ROSS C.A.;  
RT "Inositol 1,4,5-trisphosphate receptors: distinct neuronal and nonneuronal forms derived by alternative splicing differ in phosphorylation.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 88:2951-2955(1991).  
RC [-]  
CC FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
CC THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
CC [-]  
CC SUBUNIT: HOMOTETRAMER.  
CC [-]  
CC SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC RETICULUM.  
CC [-]  
CC TISSUE SPECIFICITY: THE LONGER ISOFORM IS FOUND IN THE BRAIN WHILE THE SHORTER ISOFORM IS FOUND IN THE FETAL BRAIN AND PERIPHERAL TISSUES.  
CC [-]  
CC PTM: PHOSPHORYLATED BY CYCLIC-AMP KINASE. PHOSPHORYLATION PREVENTS THE LIGAND-INDUCED OPENING OF THE CALCIUM CHANNELS.  
CC [-]  
CC CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE RECEPTOR, MOST PROBABLY BY INTERACTING WITH A DISTINCT CALCIUM-BINDING PROTEIN WHICH THEN INHIBITS THE RECEPTOR.  
CC [-]  
CC ALTERNATIVE PRODUCTS: TWO FORMS, A LONG (NEURONAL) AND A SHORT (NONNEURONAL) FORM, ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.  
CC [-]  
CC SIMILARITY: TO RYANODINE RECEPTOR.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J05510; G204675; -;  
CC EMBL; J05510; G204674; -;  
CC EMBL; M64699; G204967; -;  
CC EMBL; M64698; G204969; -;

```

QY 1 IKTKCKFLKK 10
DB 67 IKKKCKLVKK 76

RESULT 6
NBL4_MOUSE STANDARD; PRT; 554 AA.
AC P52963;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE NBL4 PROTEIN.
GN EPB4.1L4.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95074267.
RA TAKUCHI K., KAWASHIMA A., NAGAFUCHI A., TSUKITA S.;
RT "Structural diversity of band 4.1 superfamily members.";
RL J. CELL SCI. 107:1921-1928(1994).
CC -!- TISSUE SPECIFICITY: BRAIN, HEART, LUNG, LIVER AND SPLEEN. NOT
CC DETECTED IN THYMUS AND KIDNEY.
CC -!- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,
CC RADIXIN, AND TALIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D28818; G466548;
DR MGD; MGI:103007; EPB4.1L4.
DR PROSITE; PS00660; BAND_41.1; 1.
DR PROSITE; PS00661; BAND_41.2; 1.
DR PROSITE; PS50057; BAND_41.3; 1.
DR PFAM; PF00373; BAND_41.1.
KW STRUCTURAL PROTEIN; CYTOSKELETON.
FT DOMAIN 70 230 BAND 4.1-LIKE DOMAIN.
SQ SEQUENCE 554 AA; 64082 MW; 2057D0C3 CRC32;

Query Match 59.7%; Score 37; DB 1; Length 554;
Best Local Similarity 70.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KTKCKFLKK 11
DB 278 KTKCKHLWKC 287

RESULT 7
AEG1_MOUSE STANDARD; PRT; 244 AA.
AC Q03401;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE SPERM-COATING GLYCOPROTEIN 1 PRECURSOR (SCP 1) (ACIDIC EPIDIDYMAL
DE GLYCOPROTEIN 1) (CYSTEINE-RICH SECRETORY PROTEIN-1) (CRISP-1).
GN AEG1 OR AEG-1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-SUBMANDIBULAR GLAND;

```

```

EX MEDLINE; 93246016.
RA MIZUKI N., KASAHARA M.;
RT "Mouse submandibular glands express an androgen-regulated transcript
RT encoding an acidic epididymal glycoprotein-like molecule.";
RL MOL. CELL. ENDOCRINOL. 89:25-32(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-SUBMANDIBULAR GLAND;
RX MEDLINE; 93307144.
RA HAENDLER B., KRATZSCHMAR J., THEURING F., SCHLEUNING W.D.;
RT "Transcripts for cysteine-rich secretory protein-1 (CRISP-1; DE/REG)
RT and the novel related Crisp-3 are expressed under androgen control in
RT the mouse salivary gland.";
RL ENDOCRINOLOGY 133:192-198(1993).
CC -!- FUNCTION: THIS PROTEIN IS SUPPOSED TO HELP SPERMATOZOA UNDERGO
CC FUNCTIONAL MATURATION WHILE THEY MOVE FROM THE TESTIS TO THE
CC DUCTUS DEFERENS.
CC -!- SUBCELLULAR LOCATION: STORED IN SECRETORY GRANULES OF GRANULAR
CC CONVOLUTED TUBULES CELLS.
CC -!- TISSUE SPECIFICITY: MAINLY FOUND IN THE CAUDA EPIDIDYMIS WHERE IT
CC IS SYNTHESIZED BY THE PRINCIPAL CELLS AND SECRETED INTO THE LUMEN.
CC BINDS TO THE HEADS OF SPERMATOZOA. ALSO EXPRESSED IN THE
CC SUBMANDIBULAR GLAND.
CC -!- DEVELOPMENTAL STAGE: EXPONENTIAL INCREASE BETWEEN DAYS 25 AND 30
CC AFTER BIRTH.
CC -!- INDUCTION: THIS PROTEIN IS ANDROGEN-DEPENDENT.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SCI4 AND PLANTS PR-1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92849; G191738;
DR EMBL; L05559; G309191;
DR PIR; A49202; A49202.
DR MGD; MGI:102553; AEG1.
DR PROSITE; PS01009; SCP_AG5_PRL_SC7_1; 1.
DR PROSITE; PS01010; SCP_AG5_PRL_SC7_2; 1.
DR PFAM; PF00188; SCP_1.
KW SPERM; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 19
FT CHAIN 20 244 SPERM-COATING GLYCOPROTEIN 1.
FT CARBOHYD 145 145 POTENTIAL.
SQ SEQUENCE 244 AA; 27679 MW; BFBD86F2 CRC32;

Query Match 58.1%; Score 36; DB 1; Length 244;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 TKCKFLKK 10
DB 213 TKCKYLKK 220

RESULT 8
NEK2_HUMAN STANDARD; PRT; 445 AA.
AC P51955;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE SERINE/THREONINE-PROTEIN KINASE NEK2 (EC 2.7.1.-) (NIMA-RELATED
DE PROTEIN KINASE 2) (NIMA-LIKE PROTEIN KINASE 1) (HSPK 21).
GN NEK2 OR NLK1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

```

RN SEQUENCE FROM N.A.  
 RC TISSUE-T-CELL, PLACENTA, AND NASOPHARYNX;  
 RA MEDLINE: 94368699.  
 RX SCHULTZ S.J., FRY A.M., SUETTERLIN C., RIED T., NIGG E.A.;  
 RT "Cell cycle-dependent expression of Nek2, a novel human protein  
 kinase related to the NIMA mitotic regulator of Aspergillus  
 nidulans.";  
 RL CELL GROWTH DIFFER. 5:625-635(1994).  
 RN SEQUENCE FROM N.A.  
 RP LU K.P., HUNTER T.;  
 RA SUBMITTED (JUL-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE OF 83-203 FROM N.A.  
 RX MEDLINE: 94100173.  
 RA SCHULTZ S.J., NIGG E.A.;  
 RT "Identification of 21 novel human protein kinases, including 3  
 members of a family related to the cell cycle regulator nima of  
 Aspergillus nidulans.";  
 RL CELL GROWTH DIFFER. 4:821-830(1993).  
 CC -!- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION.  
 CC MAY HAVE A ROLE AT THE G2-M TRANSITION.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- DEVELOPMENTAL STAGE: ACCUMULATES THROUGHOUT S PHASE AND SHOWS  
 CC MAXIMAL LEVELS IN LATE G2. THIS EXPRESSION PATTERN IS HIGHLY  
 CC REMINISCENT OF THAT OF A AND B CYCLINS.  
 CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-  
 CC PROTEIN KINASES. BELONGS TO THE NIMA SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 229066; G479171; -;  
 DR EMBL: U11050; G507875; -;  
 DR EMBL: 225425; G405729; -;  
 DR PROSITE: P500107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE: P500108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE: P550011; PROTEIN\_KINASE\_DOM; 1.  
 DR PFAM: PF00069; pkinase; 1.  
 DR HSP: P24941; 1A01  
 KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING; MITOSIS;  
 KW NUCLEAR PROTEIN; PHOSPHORYLATION; CELL CYCLE; CELL DIVISION.  
 FT DOMAIN 8 271 PROTEIN KINASE.  
 FT NP\_BIND 14 22 ATP (BY SIMILARITY).  
 FT BINDING 37 37 ATP (BY SIMILARITY).  
 FT ACT\_SITE 141 141 BY SIMILARITY.  
 FT MOD\_RES 179 179 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CONFLICT 84 85 IV -> LY (IN REF. 2).  
 SQ SEQUENCE 445 AA; 51763 MW; 86198471 CRC32;

Query Match 58.1%; Score 36; DB 1; Length 445;  
 Best Local Similarity 77.8%; Pred. No. 34;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKCKFLKK 10

Db 407 KSKCKDLKK 415

RESULT 9

ID SMC2\_YEAST STANDARD; PRT: 1170 AA.  
 AC P38969;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE CHROMOSOME SEGREGATION PROTEIN SMC2 (DA-BOX PROTEIN SMC2).  
 GN SMC2 OR YFRO31C.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 95212908.  
 RA STRUNNIKOV A.V., HOGAN E., KOSHLAND D.;  
 RT "SMC2, a Saccharomyces cerevisiae gene essential for chromosome  
 segregation and condensation, defines a subgroup within the SMC  
 family.";  
 RL GENES DEV. 9:587-599(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-S288C / AB972;  
 RA MEDLINE: 95400292.  
 RA MURAKAMI Y., NAITOU M., HAGIWARA H., SHIBATA T., OZAWA M.,  
 RA SASANUMA S.-I., SASANUMA M., TSUCHIYA Y., SOEDA E., YOKOYAMA K.,  
 RA YAMAZAKI M., TASHIRO H., EKI T.;  
 RT "Analysis of the nucleotide sequence of chromosome VI from  
 Saccharomyces cerevisiae.";  
 RL NAT. GENET. 10:261-268(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-S288C / AB972;  
 RA MEDLINE: 96287654.  
 RA EKI T., NAITOU M., HAGIWARA H., ABE M., OZAWA M., SASANUMA S.-I.,  
 RA SASANUMA M., TSUCHIYA Y., SHIBATA T., WTANABE K., ONO A.,  
 RA YAMAZAKI M.-I., HANAOKA F., MURAKAMI Y.;  
 RT "Fifteen open reading frames in a 30.8 kb region of the right arm of  
 chromosome VI from Saccharomyces cerevisiae.";  
 RL YEAST 12:177-190(1996).  
 CC -!- FUNCTION: INVOLVED IN CHROMOSOME SEGREGATION IN MITOSIS. COULD BE  
 CC PART OF A CHROMOSOME CONDENSATION MOTOR.  
 CC -!- SUBUNIT: HOMODIMER OR HETERODIMER WITH SMC1 OR OLIGOMERS.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS  
 CC FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.  
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U05820; G468040; -;  
 DR EMBL: D50617; D1009911; -;  
 DR EMBL: D44602; D1008633; -;  
 DR PIR: S48530; S48530.  
 DR SGD: L0001927; SMC2.  
 DR HSP: P02633; 1B0D.  
 KW MITOSIS; ATP-BINDING; COILED COIL; NUCLEAR PROTEIN.  
 FT NP\_BIND 32 39 ATP (POTENTIAL).  
 FT DOMAIN 172 469 COILED COIL (POTENTIAL).  
 FT DOMAIN 678 1027 COILED COIL (POTENTIAL).  
 FT DOMAIN 1093 1119 ALA/ASP-RICH (DA-BOX).  
 SQ SEQUENCE 1170 AA; 133927 MW; C8865AC0 CRC32;

Query Match 58.1%; Score 36; DB 1; Length 1170;  
 Best Local Similarity 72.7%; Pred. No. 76;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IKTKCKFLKK 11

Db 761 IKTKQMSLKK 771

RESULT 10

POL2\_DROME  
ID POL2\_DROME STANDARD; PRT; 1059 AA.  
AC P20825;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE RETROVIRUS-RELATED POL POLYPROTEIN FROM TRANSPOSON 297 [CONTAINS:  
DE PROTEASE (EC 3.4.23.-); REVERSE TRANSCRIPTASE (EC 2.7.7.49);  
DE ENDONUCLEASE].  
GN POL.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN SEQUENCE FROM N.A.  
RP MEDLINE; 86108354.  
RX INOUE S., YUKI S., SAIGO K.:  
RT "Complete nucleotide sequence and genome organization of a Drosophila  
RT transposable genetic element, 297.";  
RL EUR. J. BIOCHEM. 154:417-425(1986).  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY U22.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X03431; NOT\_ANNOTATED\_CDS.  
DR PIR; B24872; B24872.  
DR FLYBASE; FBgn0000005; 297.  
DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
DR PFAM; PF00077; rvp; 1.  
DR PFAM; PF00078; rvt; 1.  
DR HSSP; P04585; IRTI.  
KW HYDROLASE; ASPARTYL PROTEASE; RNA-DIRECTED DNA POLYMERASE;  
KW ENDONUCLEASE; TRANSFERASE; POLYPROTEIN; TRANSPOSABLE ELEMENT.  
FT ACT\_SITE 30 30 PROTEASE (BY SIMILARITY).  
SQ SEQUENCE 1059 AA; 123310 MW; 78A65BF9 CRC32;  
  
Query Match 56.5%; Score 35; DB 1; Length 1059;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4 KCKFLKK 10  
II:IIII  
Db 399 KCEFLKK 405  
  
RESULT 11  
FBF2\_CAEEL STANDARD; PRT; 632 AA.  
AC Q09312;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE FBF-2 PROTEIN.  
GN FBF-2 OR F21H12.5.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RP FAVELLO T.  
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: C.ELEGANS F54C9.8, YEAST YGL014W, S.POMBE SPAC4G8.03C  
CC AND D.MELANOGASTER PUMILLO.  
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U23176; G726408;  
DR WORMPEP; F21H12.5; CE01916.  
DR PFAM; PF00806; PUF; 7.  
SQ SEQUENCE 632 AA; 71805 MW; 303F11BB CRC32;  
  
Query Match 54.8%; Score 34; DB 1; Length 632;  
Best Local Similarity 66.7%; Pred. No. 98;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 KTKCKFLKK 10  
II:IIII  
Db 201 KTCQFLEK 209  
  
RESULT 12  
MDHG\_SOYBN STANDARD; PRT; 353 AA.  
AC P37228;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE MALATE DEHYDROGENASE, GLYOXYMAL PRECURSOR (EC 1.1.1.37).  
OS GLYCINE MAX (SOYBEAN).  
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMERYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC FABALES; FABACEAE; PAPILIONOIDEAE; GLYCINE.  
RN [1]  
RN SEQUENCE OF 4-353 FROM N.A.  
RC STRAIN=CV. MAPLE ARROW; TISSUE=COTYLEDON;  
RX MEDLINE; 96081525.  
RA GUEX N., HENRY H., FLACH J., RICHTER H., WIDMER F.;  
RT "Glyoxysomal malate dehydrogenase and malate synthase from soybean  
RT cotyledons (Glycine max L.): enzyme association, antibody production  
RT and cDNA cloning.";  
RL PLANTA 197:369-375(1995).  
RN [2]  
RN SEQUENCE OF 1-3 FROM N.A.  
RC STRAIN=CV. MAPLE ARROW; TISSUE=COTYLEDON;  
RA RICHTER H., GUEX N.;  
RL SUBMITTED (JAN-1996) TO THE SWISS-PROT DATA BANK.  
RN [3]  
RN 3D-STRUCTURE MODELLING.  
RA GUEX N., WIDMER F., GAILLARD P., CARRUPT P.-A.;  
RT "3D modeling of soybean glyoxysomal malate dehydrogenase (gMDH).";  
RL (IN) WERMUTH C.G. (EDS.);  
RL TRENDS IN QSAR AND MOLECULAR MODELLING 92, PP.485-486, ESCOM SCIENCE  
RL PUBLISHERS, LEIDEN (1993).  
CC -!- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALACETATE + NADH.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- SUBCELLULAR LOCATION: GLYOXYSOMAL.  
CC -!- SIMILARITY: RELATED TO THE LACTATE DEHYDROGENASE CHAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: L01628; G169977;  
DR PROSITE; PS00068; MDH; 1.  
DR PFAM; PF00056; ldh; 1.  
DR HSSP; P00346; lMLD.

KW OXIDOREDUCTASE; TRICARBOXYLIC ACID CYCLE; GLYOXYISOME; NAD;  
 KW GLYOXYLATE BYPASS; TRANSIT PEPTIDE.  
 FT TRANSIT 1 33 MICROBODY (POTENTIAL).  
 FT CHAIN 34 353 GLYOXYISOMAL MALATE DEHYDROGENASE.  
 FT ACT\_SITE 190 190 PROTON-RELAY (BY SIMILARITY).  
 FT BINDING 193 193 SUBSTRATE CARBOXYL GROUP (BY SIMILARITY).  
 FT ACT\_SITE 217 217 PROTON-RELAY (BY SIMILARITY).  
 SQ SEQUENCE 353 AA; 37394 MW; C52992B4 CRC32;  
 Query Match 54.8%; Score 34; DB 1; Length 353;  
 Best Local Similarity 45.5%; Pred. No. 60;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 IKTKCKFLKKC 11  
 :|||: :||  
 Db 138 VNTLCEAIAC 148  
 :|||: :||  
 RESULT 13  
 MUTS\_AQUAE  
 ID MUTS\_AQUAE STANDARD; PRT; 859 AA.  
 AC O66652;  
 DT 15-DEC-1998 (REL. 37, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DE DNA MISMATCH REPAIR PROTEIN MUTS.  
 GN MUTS OR MUTS1.  
 OS AQUIFEX AEOLICUS.  
 OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE: 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus";  
 RL NATURE 392:353-358(1998).  
 CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES  
 CC IN DNA. IT IS POSSIBLE THAT IT CARRY OUT THE MISMATCH RECOGNITION  
 CC STEP. THIS PROTEIN HAS A WEAK ATPASE ACTIVITY (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO DNA MISMATCH REPAIR MUTS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000683; G2983001; .  
 DR PROSITE; PS00486; DNA\_MISMATCH\_REPAIR\_2; 1.  
 KW DNA REPAIR; ATP-BINDING; DNA-BINDING.  
 FT NP\_BIND 607 614 ATP (POTENTIAL).  
 SQ SEQUENCE 859 AA; 98365 MW; BDA353D6 CRC32;  
 Query Match 54.8%; Score 34; DB 1; Length 859;  
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 IKTKCKFLKKC 10  
 :|||: :||  
 Db 846 LKKKCSFSKK 855  
 :|||: :||  
 RESULT 14  
 NOF2\_DROME  
 ID NOF2\_DROME STANDARD; PRT; 403 AA.  
 AC P15297;

DT 01-APR-1990 (REL. 14, CREATED)  
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE 46 KD PROTEIN IN NOF-FB TRANSPOSABLE ELEMENT.  
 GN NOF.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OREGON-R;  
 RX MEDLINE: 89356666.  
 RA TEMPLETON N.S., POTTER S.S.;  
 RT "Complete foldback transposable elements encode a novel protein found  
 RT in Drosophila melanogaster";  
 RL EMBO J. 8:1887-1894(1989).  
 CC -!- FUNCTION: NOT KNOWN.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- SIMILARITY: TO C-TERMINAL OF NOF-FB ELEMENT PROTEIN IN  
 CC STRAIN TE146(Z).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X15469; G7964; .  
 DR PIR; S07825; S07825.  
 DR FLYBASE; FBgn0002949; NOF.  
 KW NUCLEAR PROTEIN; TRANSPOSABLE ELEMENT.  
 SQ SEQUENCE 403 AA; 46519 MW; 6A5835AF CRC32;  
 Query Match 54.8%; Score 34; DB 1; Length 403;  
 Best Local Similarity 60.0%; Pred. No. 67;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 KTKCKFLKKC 11  
 :|||: :|||  
 Db 83 KPRGKYLKKC 92  
 :|||: :|||  
 RESULT 15  
 NOF\_DROME  
 ID NOF\_DROME STANDARD; PRT; 984 AA.  
 AC P16320;  
 DT 01-AUG-1990 (REL. 15, CREATED)  
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE 112 KD PROTEIN IN NOF-FB TRANSPOSABLE ELEMENT.  
 GN NOF.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TE146(Z);  
 RX MEDLINE: 91060072.  
 RA HARDEN N., ASHBURNER M.;  
 RT "Characterization of the FB-NOF transposable element of Drosophila  
 RT melanogaster";  
 RL GENETICS 126:387-400(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA HARDEN N.;  
 RL THESIS (1989), UNIVERSITY OF CAMBRIDGE, UNITED KINGDOM.  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE TRANSDUCTION OF NOF-FB AND OTHER  
 CC FB ELEMENTS.

```
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- SIMILARITY: TO 71 KD AND 46 KD PROTEINS OF NOF-FB ELEMENT IN
CC STRAIN OREGON-R.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X51937; G829181; -
CC PIR: S14382; S14382.
CC PIR: S14383; S14383.
CC FLYBASE: FBgn0002949; NOF.
CC NUCLEAR PROTEIN; TRANSPOSABLE ELEMENT.
CC NON_TER 1
CC SEQUENCE 984 AA; 112748 MW; 234C3AE5 CRC32;
CC -----
Query Match
Best Local Similarity 54.8%; Score 34; DB 1; Length 984;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 2 KTKCKFLAKC 11
Db 675 KPRGKYLAKC 684
| : | : | : | : |
| : | : | : | : |

Search completed: September 7, 1999, 23:59:15
Job time: 527 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:46 ; Search time 160.2 Seconds  
(without alignments)  
4.226 Million cell updates/sec

Title: US-09-124-280A-35

Perfect score: 62

Sequence: 1 IKTKCKFLKKC 11

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	62.9	971	5	O45785
2	39	62.9	274	10	O82763
3	39	62.9	125	12	O68540
4	38.5	62.1	239	12	O35741
5	38	61.3	347	2	O51533
6	38	61.3	2713	4	O14660
7	38	61.3	2695	4	O14643
8	38	61.3	422	12	O9VW4
9	37	59.7	382	2	O85003
10	37	59.7	217	2	O9ZAC5
11	37	59.7	1121	3	O12734
12	37	59.7	504	11	O60565
13	36.5	58.9	1070	3	P78734
14	36.5	58.9	472	5	O25035
15	36	58.1	170	1	O59556
16	36	58.1	492	2	O25029
17	36	58.1	491	2	O9ZM18
18	36	58.1	583	5	O27472
19	36	58.1	548	5	O19496
20	36	58.1	200	5	O23170
21	36	58.1	284	10	O39862
22	36	58.1	319	10	O40780
23	36	58.1	318	10	O40781
24	36	58.1	302	10	O40782
25	36	58.1	171	10	O43441
26	36	58.1	154	12	O70732
27	36	58.1	154	12	O70734
28	36	58.1	471	12	O84697
29	36	58.1	1009	12	O9VW4

30	35	56.5	379	2	O68797	O68797 campylobact
31	35	56.5	135	2	O32236	O32236 bacillus su
32	35	56.5	449	2	O92M35	O92M35 helicobacte
33	35	56.5	60	2	O92383	O92383 escherichia
34	35	56.5	1539	4	O15078	O15078 homo sapien
35	35	56.5	2671	4	O14649	O14649 homo sapien
36	35	56.5	2671	4	O14573	O14573 homo sapien
37	35	56.5	151	5	O15655	O15655 plasmodium
38	35	56.5	294	5	O18972	O18972 caenorhabdi
39	35	56.5	125	5	P91462	P91462 caenorhabdi
40	35	56.5	703	5	O44403	O44403 caenorhabdi
41	35	56.5	486	5	O01822	O01822 caenorhabdi
42	35	56.5	743	5	O17305	O17305 caenorhabdi
43	35	56.5	1270	5	O62462	O62462 caenorhabdi
44	35	56.5	1122	5	O96247	O96247 plasmodium
45	35	56.5	1139	5	O97242	O97242 plasmodium

#### ALIGNMENTS

##### RESULT 1

O45785 PRELIMINARY: PRT: 971 AA.  
AC O45785; O45816;  
DT 01-JUN-1998 (Tremblrel. 06, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DE 01-JAN-1999 (Tremblrel. 09, Last annotation update)  
DE T12D8.9 PROTEIN.  
GN T12D8.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoides; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER A.;  
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 94150718.  
RA WILSON R.; AINSCOUGH R.; ANDERSON K.; BAYNES C.; BERKS M.;  
RA BONFIELD J.; BURTON J.; CONNELL M.; COPSEY T.; COOPER J.; COULSON A.;  
RA CRAXTON M.; DEAR S.; DU Z.; DURBIN R.; FAVELLO A.; FULTON L.;  
RA GARDNER A.; GREEN P.; HAWKINS T.; HILLIER L.; JIER M.; JOHNSTON L.;  
RA JONES M.; KERSHAW J.; KIRSTEN J.; LAISTER N.; LATREILLE P.;  
RA LIGHTNING J.; LLOYD C.; MCMURRAY A.; MORTIMORE B.; O'CALLAGHAN M.;  
RA PARSONS J.; PERCY C.; RIEKEN L.; ROOPRA A.; SAUNDERS D.; SHOWNKEEN R.;  
RA SMALDON N.; SMITH A.; SONNHAMMER E.; STADEN R.; SULSTON J.;  
RA THIERRY-MIEG J.; THOMAS K.; VAUDIN M.; VAUGHAN K.; WATERSTON R.;  
RA WATSON A.; WEINSTOCK L.; WILKINSON-SPROAT J.; WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA MCMURRAY A.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: 283241; CAB05819.1; -.  
DR EMBL: 281120; CAB05819.1; JOINED.  
DR EMBL: 281120; CAB03350.1; -.  
DR EMBL: 283241; CAB03350.1; JOINED.  
SQ SEQUENCE 971 AA; 113043 MW; 5FF79B6D CRC32;

Query Match 62.9%; Score 39; DB 5; Length 971;  
Best Local Similarity 54.5%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IKTKCKFLKKC 11

DB 340 LKTLCKYRKSC 350

```

RESULT 2
O82763 ID O82763 PRELIMINARY; PRT; 274 AA.
AC O82763;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE HOMEBOX PROTEIN, HAT9.
GN T30L20.6 OR T20K9.1.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eudicotyledons; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,
RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;
PT "Arabidopsis thaliana chromosome II BAC T30L20 genomic sequence.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,
RA SYKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,
RA SOMERVILLE C.R., VENTER J.C.;
RT "Arabidopsis thaliana chromosome II BAC T20K9 genomic sequence.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DE EMBL; AC005617; AAC63591.1; -
DR EMBL; AC004786; AAC32427.1; -
DR PFAM; PF00046; homeobox; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
SQ Homeobox; DNA-binding; Nuclear protein.
SQ SEQUENCE 274 AA; 29877 MW; 0797F7C1 CRC32;

Query Match 62.9%; Score 39; DB 10; Length 274;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 5 CKFLKCC 11
Db 174 CEFLKCC 180

RESULT 3
O68540 ID O68540 PRELIMINARY; PRT; 125 AA.
AC O68540;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE MOVEMENT PROTEIN V2.
GN V2.
OS Horseradish curly top virus.
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA KLUTE K.A., NADLER S.A., STENGER D.C.;
RT "Horseradish curly top virus is a distinct subgroup II geminivirus
RT species with rep and C4 genes derived from a subgroup III ancestor.";
RL J. Gen. Virol. 77:1369-1378(1996).
DR EMBL; U418907; AAB18923.1; -
SQ SEQUENCE 125 AA; 14506 MW; 580ED7E5 CRC32;

Query Match 62.9%; Score 39; DB 12; Length 125;
Best Local Similarity 54.5%; Pred. No. 6.1;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RESULT 4
O55741 ID O55741 PRELIMINARY; PRT; 239 AA.
AC O55741;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TRENBLrel. 06, Last annotation update)
DE HYPOHETICAL 28.0 KD PROTEIN.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA BAHU U., TIDONA C.A., DARAI G.;
RL Virus Genes 0:0-0(1997).
DR EMBL; AF003534; AAB94452.1; -
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 27979 MW; 7B15C915 CRC32;

Query Match 62.1%; Score 38.5; DB 12; Length 239;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

Oy 2 KTK-CKFLKCC 11
Db 83 KTRMCKYVKKC 93

RESULT 5
O51533 ID O51533 PRELIMINARY; PRT; 347 AA.
AC O51533;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE FEMA PROTEIN (FEMA).
GN BB0586.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., OUCKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AE001160; AAB91521.1; -
DR TIGR; BB0586; -
SQ SEQUENCE 347 AA; 41521 MW; BA5C09E8 CRC32;

Query Match 61.3%; Score 38; DB 2; Length 347;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 IKTKCKFLKK 10
Db 120 LKTRIKYVKK 129

```

```

RESULT 6
Q14660
ID Q14660 PRELIMINARY; PRT; 2713 AA.
AC Q14660;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE INOSITOL 1,4,5-TRISPHOSPHATE RECEPTOR TYPE 1.
GN ITPR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-T LYMPHOCYTE;
RX MEDLINE; 95155356.
RA HARNICK D.J., JAYARAMAN T., MA Y., MULIERI P., GO L.O., MARKS A.R.;
RT "The human type 1 inositol 1,4,5-trisphosphate receptor from T
RT lymphocytes. Structure, localization, and tyrosine phosphorylation.";
RL J. Biol. Chem. 270:2833-2840(1995).
DR EMBL; L38019; AAB04947.1; -.
DR PFAM; PF01365; RYDR_ITPR; 1.
SQ SEQUENCE 2713 AA; 308697 MW; A16B46FB CRC32;

Query Match 61.3%; Score 38; DB 4; Length 2713;
Best Local Similarity 54.5%; Pred. No. 89;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11
: | : | : | : |
Db 1318 VKAEGKFIKK 1328

RESULT 7
Q14643
ID Q14643 PRELIMINARY; PRT; 2695 AA.
AC Q14643;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HUMAN TYPE 1 INOSITOL 1,4,5-TRISPHOSPHATE RECEPTOR.
GN INSP3RI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-UTERUS, AND MYELOID;
RX MEDLINE; 95031918.
RA YAMADA N., MAKINO Y., CLARK R.A., PEARSON D.W., MATTEI M.G.,
RA GUENET J.L., OHAMA E., FUJINO I., MIYAWAKI A., FURUICHI T.,
RA MIKOSHIBA K.;
RT "Human inositol 1,4,5-trisphosphate type-1 receptor, INSP3RI:
RT structure, function, regulation of expression and chromosomal
RT localization.";
RL Biochem. J. 302:781-790(1994).
DR EMBL; D26070; BAA05065.1; -.
DR PFAM; PF01365; RYDR_ITPR; 1.
SQ SEQUENCE 2695 AA; 306769 MW; 99197877 CRC32;

Query Match 61.3%; Score 38; DB 4; Length 2695;
Best Local Similarity 54.5%; Pred. No. 89;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11
: | : | : | : |
Db 1302 VKAEGKFIKK 1312

RESULT 8
Q9YVW4

```

```

ID Q9YVW4 PRELIMINARY; PRT; 422 AA.
AC Q9YVW4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE ORF MSV128 HYPOTHETICAL PROTEIN.
GN MSV128.
OS Melanoplus sanguinipes entomopoxvirus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TUCSON;
RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.;
RT "The genome of Melanoplus sanguinipes entomopoxvirus.";
RL J. Virol. 73:533-552(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TUCSON;
RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF063866; AAC97795.1; -.
SQ SEQUENCE 422 AA; 50633 MW; 93A75459 CRC32;

Query Match 61.3%; Score 38; DB 12; Length 422;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKCKFLKK 10
: | : | : | : |
Db 290 LQKCKYLKK 299

RESULT 9
O85003
ID O85003 PRELIMINARY; PRT; 382 AA.
AC O85003;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE GLYCERO-PHOSPHOTRANSFERASE.
GN CPS23FW.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SP-264;
RX MEDLINE; 98125733.
RA COFFEY T.J., ENRIGHT M.C., DANIELS M., MORONA J.K., MORONA R.,
RA HRYNIEWICZ W., PATON J.C., SPRATT B.G.;
RT "Recombinational exchanges at the capsular polysaccharide
RT biosynthetic locus lead to frequent serotype changes among natural
RT isolates of Streptococcus pneumoniae.";
RL Mol. Microbiol. 27:73-83(1998).
DR EMBL; AF030373; AAC38753.1; -.
KW Transferase.
SQ SEQUENCE 382 AA; 44883 MW; 01B5F413 CRC32;

Query Match 59.7%; Score 37; DB 2; Length 382;
Best Local Similarity 87.5%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKCKFL 8
: | : | : | : |
Db 240 IKKCKFL 247

RESULT 10
Q9ZAC5
ID Q9ZAC5 PRELIMINARY; PRT; 217 AA.

```

AC Q9ZAC5;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE HYPOTHETICAL 25.6 KD PROTEIN.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 CC Campylobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCTC 11168;  
 RX MEDLINE; 99173742.  
 RA WOOD A.C., OLDFIELD N.J., O'DWYER C.A., KETLEY J.M.;  
 RT "Cloning, mutation and distribution of a putative lipopolysaccharide  
 biosynthesis locus in Campylobacter jejuni.";  
 RL Microbiology 145:0-0(0).  
 DR EMBL; AF001497; AAD09297.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 217 AA; 25634 MW; 9874B248 CRC32;

Query Match 59.7%; Score 37; DB 2; Length 217;  
 Best Local Similarity 72.7%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11  
 |||||  
 Db 186 IKTKCKFLKCC 196

RESULT 11  
 Q12734  
 ID Q12734 PRELIMINARY; PRT; 1121 AA.  
 AC Q12734;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last annotation update)  
 DE HYPOTHETICAL 124.8 KD PROTEIN YPR029W.  
 GN YPR029W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 RN Saccharomycetaceae; Saccharomyces.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB972;  
 RA BADCOCK K., BOWMAN S., CHURCHER C.M., PEARSON D., RAJANDREAM M.A.,  
 RA WALSH S.V., BARRELL B.G.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB972;  
 RA BADCOCK K., CHURCHER C.M.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB972;  
 RA BARRELL B., RAJANDREAM M.A.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z71255; CAA95026.1; -.  
 DR EMBL; Z49274; CAA89284.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1121 AA; 124847 MW; 7997D462 CRC32;

Query Match 59.7%; Score 37; DB 3; Length 1121;  
 Best Local Similarity 63.6%; Pred. No. 69;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 11  
 |||||  
 Db 773 VKTKLKPCKYC 783

RESULT 12  
 Q60565  
 ID Q60565 PRELIMINARY; PRT; 504 AA.  
 AC Q60565;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE POTASSIUM CHANNEL KV8.1.  
 OS Mesocricetus auratus (Golden hamster).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96283631.  
 RA HUGNOT J.P., SALINAS M., LESAGE F., GUILLEMARE E., DE WEILLE J.,  
 RA HEURTEAUX C., MATTI M.G., LAZDUNSKI M.;  
 RT "Kv8.1, a new neuronal potassium channel subunit with specific  
 inhibitory properties towards Shab and Shaw channels.";  
 RL EMBL; U62810; AAC52727.1; -.  
 DR PFAM; PF00520; ion\_trans; 1.  
 KW Ionic channel.  
 SQ SEQUENCE 504 AA; 56725 MW; 92EIF6D9 CRC32;

Query Match 59.7%; Score 37; DB 11; Length 504;  
 Best Local Similarity 50.0%; Pred. No. 38;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKCKFLKCC 10  
 :||:|:|:  
 Db 268 VKDCRFLKCC 277

RESULT 13  
 P78734  
 ID P78734 PRELIMINARY; PRT; 1070 AA.  
 AC P78734;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
 DE MAJOR SURFACE GLYCOPROTEIN.  
 OS Pneumocystis carinii.  
 CC Eukaryota; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA EDMAN J.C., HATTON T.W., KOVACS J.A.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U83323; AAB41323.1; -.  
 SQ SEQUENCE 1070 AA; 120810 MW; 7B3FD4DE CRC32;

Query Match 58.9%; Score 36.5; DB 3; Length 1070;  
 Best Local Similarity 66.7%; Pred. No. 81;  
 Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 IKTKC-KFLKCC 11  
 :|||:|:  
 Db 314 VRECKCKALKCC 325

RESULT 14  
 Q25035  
 ID Q25035 PRELIMINARY; PRT; 472 AA.  
 AC Q25035;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE GABA A/GLYCINE RECEPTOR SUBUNIT.  
 GN HGI.  
 OS Haemophilus contortus.  
 CC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Strongylida;

OC Trichostrongyloidea; Trichostrongylidae; Haemonchinae; Haemonchus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95041083.  
RA LAUGHTON D.L., AMAR M., THOMAS P., TOWNER P., HARRIS P., LUNT G.C.,  
RA WOLSTENHOLME A.J.;  
RT "Cloning of a putative inhibitory amino acid receptor subunit from  
RT the parasitic nematode Haemonchus contortus.";  
RL Recept. Channels 2:155-163(1994).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
DR EMBL; X73584; CAA51991.1; .  
DR PFAM; PF00065; neur\_chan; 2.  
DR PROSITE; PS00236; NEUROTR\_ION\_CHANNEL; 1.  
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane.  
FT VARIANT 231 231 G -> D.  
FT VARIANT 329 329 S -> A.  
SQ SEQUENCE 472 AA; 54732 MW; 7B8A2213 CRC32;

Query Match 58.9%; Score 36.5; DB 5; Length 472;  
Best Local Similarity 81.8%; Pred. No. 44;  
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 IRTKC-KFLKK 10  
| | | | | | | |  
DB 160 IRTKCLMFLKK 170

RESULT 15  
ID O59556 PRELIMINARY; PRT; 170 AA.  
AC O59556;  
DT 01-AUG-1998 (TRENBLrel. 07, Created)  
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)  
DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
DE 170AA LONG HYPOTHETICAL PROTEIN.  
GN PH1864.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OT3;  
RX MEDLINE; 98344137.  
RA KAWARABAYASHI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOVAMA A., NAGAI Y.,  
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
RA KIKUCHI H.;  
RT "Complete sequence and gene organization of the genome of a  
RT hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res 5:55-76(1998).  
DR EMBL; AF000007; BAA30986.1; .  
SQ SEQUENCE 170 AA; 19813 MW; F9DEC0A1 CRC32;

Query Match 58.1%; Score 36; DB 1; Length 170;  
Best Local Similarity 54.5%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 IRTKCKFLKK 11  
: | | | | |  
DB 15 VKFKCKFLDC 25

Search completed: September 7, 1999, 22:47:46  
Job time: 7966 sec

**THIS PAGE BLANK (b)(7)(C)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:06 ; Search time 135.78 Seconds  
(without alignments)  
1.744 Million cell updates/sec

Title: US-09-124-280A-36

Perfect score: 49

Sequence: 1 IKTKFLKKT 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	49	100.0	1 R33528	Peptide for treati
2	49	100.0	1 R33528	Endotoxin lipid A
3	49	100.0	1 R57882	Lipid A minimum bi
4	49	100.0	1 W21624	Antibiotic potenti
5	49	100.0	1 W62433	Human neutrophil g
6	33.5	68.4	1 R33531	Peptide for treati
7	33.5	68.4	1 R33531	Endotoxin lipid A
8	33.5	68.4	1 W21623	Antibiotic potenti
9	33	67.3	1 W77567	Staphylococcus aur
10	32	65.3	1 W59045	Human MNTF1-F3 pro
11	31	63.3	1 R05435	Feline herpes viru
12	31	63.3	1 R23996	B-HDAG-T recombin
13	31	63.3	1 R11507	HDV delta antigen
14	31	63.3	1 P82521	Hepatitis D Virus
15	31	63.3	1 R64784	Amphipathic peptid
16	31	63.3	1 R64785	Amphipathic peptid
17	31	63.3	1 R64786	Amphipathic peptid
18	31	63.3	1 R64790	Amphipathic peptid
19	31	63.3	1 R64791	Amphipathic peptid
20	31	63.3	1 R64792	Amphipathic peptid
21	31	63.3	1 R74705	Tryptic digestion
22	31	63.3	1 R74706	Tryptic digestion
23	31	63.3	1 R74707	Tryptic digestion
24	31	63.3	1 R74711	Tryptic digestion
25	31	63.3	1 R74712	Tryptic digestion
26	31	63.3	1 R74713	Tryptic digestion
27	31	63.3	1 R77056	Synthetic anti-neo
28	31	63.3	1 R77057	Synthetic anti-neo
29	31	63.3	1 R77058	Synthetic anti-neo
30	31	63.3	1 R77062	Synthetic anti-neo
31	31	63.3	1 R77063	Synthetic anti-neo
32	31	63.3	1 R77064	Synthetic anti-neo
33	31	63.3	1 R84150	Peptide enhancer o
34	31	63.3	1 R84142	Peptide enhancer o
35	31	63.3	1 R84143	Peptide enhancer o
36	31	63.3	1 R84144	Peptide enhancer o
37	31	63.3	1 R84148	Peptide enhancer o
38	31	63.3	1 R84149	Peptide enhancer o
39	31	63.3	1 R92386	Lytic peptide used
40	31	63.3	1 R92387	Lytic peptide used
41	31	63.3	1 R92388	Lytic peptide used
42	31	63.3	1 R92392	Lytic peptide used
43	31	63.3	1 R92393	Lytic peptide used

44 31 63.3 37 1 R92394  
45 31 63.3 968 1 W21721  
Lytic peptide used  
SH2-A. Detection o

ALIGNMENTS

RESULT 1  
R33528  
ID R33528 standard; peptide; 10 AA.  
AC R33528;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 9; Page 32; 39pp; English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKFLKKT 10  
DB 1 IKTKFLKKT 10  
|||||

RESULT 2  
R39293  
ID R39293 standard; peptide; 10 AA.  
AC R39293;  
DT 22-DEC-1993 (first entry)  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW polymyxin B; PWB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
OS Synthetic.  
PN WO9314115-A.  
PD 22-JUL-1993.  
PF 14-MAY-1992; E01060.  
PR 16-JAN-1992; US-819893.  
PA (PORR/) PORRO M.

PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 9; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKKT 10  
 DB 1 IKTKKFLKKT 10

RESULT 3  
 R57882  
 ID R57882 standard; peptide; 10 AA.  
 AC R57882;  
 DT 28-MAR-1995 (first entry)  
 DE Lipid A minimum binding site.  
 KW Binding site; CDR; complementarity determining region; immunoglobulin;  
 KW heavy; light; primer extension; PCR; amplif; fibronectin; vitronectin;  
 KW RGD-dependent; integrin ligand; von Willebrand factor; EBV; gp350/220;  
 KW envelope glycoprotein; HIV; gp120; reovirus; hemagglutinin; insulin;  
 KW cellular receptor; CR2; CD4; hormone; thyroid stimulating hormone; TSH;  
 KW transferrin; apolipoprotein; apo E; apo A1; MHC; class I; class II;  
 KW non-RGD-dependent; vitronectin receptor; alpha-v; beta-3; modulation;  
 KW anti-gp11b/IIIa; monoclonal antibody; MAB; platelet adhesion; cancer;  
 KW coagulation; inflammation; anti-vitronectin; tumour cell adhesion;  
 KW migration.  
 OS Homo sapiens.  
 PN WO9418221-A.  
 PD 18-AUG-1994.  
 PF 02-FEB-1994; U01258.  
 PR 02-FEB-1993; US-012566.  
 PR 28-JUN-1993; US-084542.  
 PA (SCRI ) SCRIPPS RES INST.  
 PI Barbas CF, Lerner RA;  
 DR WPI; 94-279675/34.  
 PT Production of binding sites within CDR regions of immunoglobulins  
 PT - displayed on the surface of filamentous phage particles, for  
 PT inhibiting platelet aggregation and vitronectin binding  
 PS Disclosure; Page 27; 207pp; English.  
 CC The sequences given in R57837-84 are binding sites which were used in  
 CC the method of the invention for producing a polypeptide having a  
 CC binding site capable of binding a preselected agent. Nucleotide  
 CC sequences encoding these binding site peptides were introduced into  
 CC a CDR region of a nucleic acid encoding an immunoglobulin heavy (H)  
 CC or light (L) chain, by amplifying the CDR region by primer extension.  
 CC Preferred binding sites are derived from the RGD-dependent integrin  
 CC ligands, eg. fibronectin, vitronectin, von Willebrand factor, from  
 CC the envelope glycoprotein from viruses such as HIV gp120, EBV gp350/  
 CC 220, reovirus hemagglutinin, from cellular receptors such as CR2 or  
 CC CD4, from protein hormones such as thyroid stimulating hormone (TSH),  
 CC insulin, transferrin, from apolipoproteins such as apo E and apo A1,  
 CC from immunoglobulin CDRs and from MHC class I or II proteins. Non-RGD-  
 CC dependent integrin binding sites were selected for the affinity to bind  
 CC vitronectin receptor alpha-v, beta-3. An anti-gp11b/IIIa monoclonal

CC antibody (MAB) produced in this way can be used to modulate platelet  
 CC adhesion in the treatment of coagulation and some inflammatory responses.  
 CC An anti-vitronectin MAB can be used in the treatment of cancer by  
 CC blocking tumour cell adhesion and migration. This sequence represents  
 CC the lipid A binding site.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKKT 10  
 DB 1 IKTKKFLKKT 10

RESULT 4  
 W21624  
 ID W21624 standard; peptide; 10 AA.  
 AC W21624;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #36.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PR 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 40; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKKT 10  
 DB 1 IKTKKFLKKT 10

RESULT 5  
 W62433  
 ID W62433 standard; peptide; 10 AA.  
 AC W62433;  
 DT 01-OCT-1998 (first entry)  
 DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:48.  
 KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;  
 KW bactericidal permeability increasing factor; B/PI; neutralising;  
 KW human neutrophil granule bactericidal protein.  
 OS Homo sapiens.  
 PN US5786324-A.  
 PD 28-JUL-1998.  
 PR 24-MAR-1994; 218026.  
 PR 24-MAR-1994; US-218026.  
 PA (MINU ) UNIV MINNESOTA.



PI Gray B, Haseman JR, Mayo K;  
PT WPI: 98-436578/37.  
PT Bactericidal and endotoxin-neutralising peptides - used in treating  
PT e.g. Pseudomonas species infection and in protectively coating  
PT prosthetic devices  
PS Example 1: Column 18: 46pp; English.  
CC The present invention describes bactericidal peptides (BP): (a) for  
CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)  
CC with both endotoxin neutralising activity and bactericidal activity for  
CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising  
CC activity and bactericidal activity, the peptide having at least 1 amino  
CC acid sequence of a peptide with endotoxin neutralising activity combined  
CC with at least 1 different amino acid sequence of a peptide that has  
CC bactericidal activity; and (e) a prosthetic device which has a  
CC sufficient amount of BP attached to the surface to inhibit bacterial  
CC growth. The peptides of the invention are used in treating bacterial  
CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10-7 to  
CC 10-9 M, and Escherichia coli. The peptides are also used to treat  
CC endotoxin shock. The present sequence represents a peptide derived  
CC from human neutrophil granule bactericidal protein from an example of  
CC the present invention.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKKT 10  
Db 1 IKTKKFLKKT 10

RESULT 6  
R33531  
ID R33531 standard; peptide: 11 AA.  
AC R33531;  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 5..11  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI: 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 8; Page 32; 39pp; English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 11 AA;

Query Match 68.4%; Score 33.5; DB 1; Length 11;  
Best Local Similarity 90.0%; Pred. No. 2.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKTK-KFLKK 9  
Db 1 IKTKCKFLKK 10

RESULT 7  
R39292  
ID R39292 standard; peptide: 11 AA.  
AC R39292;  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
KW quantification.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 5..11  
PN WO9314115-A.  
PD 22-JUL-1993.  
PF 14-MAY-1992; E01060.  
PR 16-JAN-1992; US-819893.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI: 93-243143/30.  
PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
PT forming non-toxic, antigenic complex, used to treat or prevent  
PT septic shock, in vaccines to detoxify blood, etc.  
PS Claim 8; Page 33; 45pp; English.  
CC The sequence is that of a peptide which binds to the lipid A  
CC component of bacterial endotoxin at the same site as polymyxin B  
CC (PMB) and with about the same affinity to produce a non-toxic  
CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
CC proteolytic degradation in serum, has no antibiotic activity and no  
CC haemolytic action. It is especially used to treat or prevent septic  
CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
CC sera, vaccines, drug solns. etc.; to control release of cytokines  
CC induced by endotoxins; for in vivo or in vitro detoxification of  
CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
CC products.  
SQ Sequence 11 AA;

Query Match 68.4%; Score 33.5; DB 1; Length 11;  
Best Local Similarity 90.0%; Pred. No. 2.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKTK-KFLKK 9  
Db 1 IKTKCKFLKK 10

RESULT 8  
W21623  
ID W21623 standard; peptide: 11 AA.  
AC W21623;  
DE 26-AUG-1997 (first entry)  
KW Antibiotic potentiating peptide #35.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 5..11  
PN WO9638163-A1.  
PR 05-DEC-1996.  
PD 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI POIRO M, Vaira M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 39; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 CC Sequence 11 AA;  
 SQ

Query Match 68.4%; Score 33.5; DB 1; Length 11;  
 Best Local Similarity 90.0%; Pred. No. 2.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 IKTK-KFLKK 9  
 DB 1 IKTRCKFLKK 10

RESULT 9  
 W77567  
 ID W77567 standard; Protein; 119 AA.  
 AC W77567;  
 DT 30-OCT-1998 (first entry)  
 DE Staphylococcus aureus protein of unknown function.  
 KW Staphylococcus aureus protein; immune response induction; eye infection;  
 KW antibody production; i-cell immune response; gastrointestinal infection;  
 KW respiratory infection; inhibitor; bacterial infection; cardiac infection;  
 KW central nervous system; kidney infection; urinary tract infection;  
 KW antimicrobial compound identification; broad spectrum antibiotic;  
 KW therapy.  
 OS Staphylococcus aureus.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 1119 /note= "residues designated X are unspecified, and  
 represented as Xaa in the specification"  
 FT EP-841394-A2.  
 PN 13-MAY-1998.  
 PD 24-SEP-1997; 307485.  
 PR 24-SEP-1996; US-027032.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Black MT, Burnham MKR, Hodgson JE, Knowles DJC,  
 PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M,  
 PI Ward JM;  
 DR WPI: 98-252940/23.  
 DR N-PSDB; V53367.  
 PT New nucleic acid sequences from Staphylococcus aureus WCHU29 -  
 PT useful in vaccines and for treatment of bacterial infections of e.g.  
 PT respiratory tract and central nervous system  
 PS Claim 11; Page 256-257; 390pp; English.  
 CC This sequence represents a Staphylococcus aureus protein of unknown  
 CC function, and is encoded by a DNA sequence of the invention.  
 CC The DNA sequences were isolated from Staphylococcus aureus WCHU29  
 CC (NCIMB 40771). Host cells containing the DNA sequences are used to  
 CC produce polypeptides or fragments. The proteins are used in the treatment  
 CC of disease, for inducing an immune response by administering them, to  
 CC produce antibody and/or T-cell immune response. Antagonists of the  
 CC proteins are used for the inhibition of bacterial polypeptides.  
 CC Conditions which may be treated include bacterial infections, especially  
 CC respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,  
 CC urinary tract, skin, bones and joints. The proteins can also be used to  
 CC identify antimicrobial compounds which are broad spectrum antibiotics,  
 CC especially useful in the treatment of H. pylori infection.

SQ Sequence 119 AA;  
 Query Match 67.3%; Score 33; DB 1; Length 119;  
 Best Local Similarity 66.7%; Pred. No. 30;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 IKTKKFLKK 9  
 DB 30 IKTSKYVK 38

RESULT 10  
 W59045  
 ID W59045 standard; Protein; 481 AA.  
 AC W59045;  
 DT 11-AUG-1998 (first entry)  
 DE Human MNTF1-F3 protein fragment.  
 KW Motoneuronotrophic factor; MNTF-1; MNTF1-F3; human; axon regeneration;  
 KW motoneurone; diagnose; treatment; disease; wound healing; scar tissue;  
 KW keloid.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 194 /note= "encoded by TTC"  
 FT Misc\_difference 248 /note= "encoded by AGC"  
 FT Misc\_difference 461 /note= "encoded by ACY"  
 PN W09813492-A2.  
 PD 02-APR-1998.  
 PF 22-SEP-1997; U17142.  
 PR 12-SEP-1997; US-928862.  
 PR 27-SEP-1996; US-026792.  
 PR 15-NOV-1996; US-751225.  
 PA (KMBI-) KM BIOTECH INC.  
 PI Chau RMW;  
 DR WPI: 98-230703/20.  
 DR N-PSDB; V11747.  
 PT Motoneuronotrophic factor MNTF1-F3 and MNTF1-F6 - useful for  
 PT motoneuron regeneration, diagnosing or treating motoneuron disease  
 PT and to accelerate wound healing without scar formation  
 PS Claim 3; Fig 2A; 78pp; English.  
 CC This sequence represents a novel human motoneuronotrophic factor, MNTF1-F3.  
 CC Such factors are used to promote regeneration of the axon of a  
 CC motoneurone, to diagnose and treat motoneurone disease in a mammal or to  
 CC accelerate wound healing whilst concomitantly minimising or inhibiting  
 CC scar tissue and/or keloid formation in an area associated with a wound.  
 CC For promoting axonal regeneration, the polypeptide is administered at a  
 CC concentration of 5 ng-50 mg, whereas for inhibiting hereditary  
 CC motoneurone disease, the dosage is 5-100 (especially 30-50) ng per kg  
 CC body weight.  
 CC Sequence 481 AA;  
 SQ

Query Match 65.3%; Score 32; DB 1; Length 481;  
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 KTKKFLKK 9  
 DB 64 KTRKFTKK 71

RESULT 11  
 R05435  
 ID R05435 standard; protein; 343 AA.  
 AC R05435;  
 DT 30-JUL-1990 (first entry)  
 DE Feline herpes virus thymidine kinase.  
 KW Thymidine kinase; herpes virus; vaccine; ss.  
 OS Feline Herpes Virus strain UC-D.  
 PN W09001547-A.

PD 22-FEB-1990.  
 PF 02-AUG-1989; U03289.  
 PR 12-JUL-1989; US-379041.  
 PA (CETU) Cetus Corp (UPJO).  
 PI Post LE, Compton T, Nunberg JH, Petrovskis EA;  
 DR WPI: 90-083506/11.  
 DR N-PSDB; Q03570.  
 PT Isolation of thymidine kinase-encoding DNA from herpes virus -  
 PT from degenerative primers, and prodn. of thymidine  
 PT kinase-negative feline herpes virus used to produce live vaccine.  
 PS Claim 8; Page 25; 37pp; English.  
 CC The recombinant TK-negative fHSV is used  
 CC for the construction of live vaccines.  
 SQ Sequence 343 AA;

Query Match 63.3%; Score 31; DB 1; Length 343;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKKFLKK 9  
 : |||||  
 DB 215 VHTKKFLTK 223

RESULT 12  
 R23996 ID R23996 standard; Protein: 310 AA.  
 AC R23996;  
 DT 09-NOV-1992 (first entry)  
 DE B-HDAG-T recombinant polypeptide.  
 KW Beta-galactosidase; linker; HDAG; pBR322 tet gene; HDV;  
 KW hepatitis delta viral antigen; hepatitis delta virus;  
 KW immunoassays.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT region 1..8  
 FT /note= "B-HDAG-T antigenic fragment"  
 FT region 9..12  
 FT /note= "beta-galactosidase fragment"  
 FT region 13..221  
 FT /note= "linker"  
 FT region 222..310  
 FT /note= "fragment encoded by pBR322 tet gene"

PN EP-485347-A.  
 PD 13-MAY-1992.  
 PF 04-NOV-1991; 830479.  
 PR 05-NOV-1990; IT-067865.  
 PA (SORI-) SORIN BIOMEDICA SPA.  
 PI Bonelli F, Boniolo A, Calogero R;  
 DR WPI: 92-161087/20.  
 DR N-PSDB; Q24686.  
 PT Compens. contg. polymeric aggregates of recombinant HDAG - useful  
 PT in diagnosis of HDV infection, e.g. by immunoassay of anti-HDAG  
 PT IgM  
 PS Disclosure; Page 4; 23pp; English.  
 CC The sequence is that of hepatitis delta antigen (HDAG)-like recombinant  
 CC polypeptide B-HDAG-T which may be isolated in multimeric form from the  
 CC cytosolic fraction of transformed bacterial cells using a simple one  
 CC step ultracentrifugation procedure. The recovered polypeptide has the  
 CC antigenic epitopes preserved and has antigenicity comparable to the  
 CC intact natural antigen. It is therefore useful in immunoassays of HDAG  
 CC or HDAG-specific antibodies present in tissue or fluid samples, as such  
 CC it can be used for the detection of hepatitis delta virus (HDV)  
 CC infection. 310 AA;  
 SQ Sequence 310 AA;

Query Match 63.3%; Score 31; DB 1; Length 310;  
 Best Local Similarity 87.5%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9

DB 48 KTKKFLKK 55  
 |||||  
 RESULT 13  
 R11507 ID R11507 standard; Protein: 55 AA.  
 AC R11507;  
 DT 14-JUN-1991 (first entry)  
 DE HDV delta antigenic epitope  
 KW Hepatitis delta virus; HDV; exacerbated liver disease.  
 OS Synthetic.  
 PN W09104341-A.  
 PD 04-APR-1991.  
 PF 21-SEP-1990; U05406.  
 PR 21-SEP-1989; US-410068.  
 PR 20-MAR-1990; US-498085.  
 PA (IMCL-) IMCLONE SYST INC.  
 PI Tackney CT, Waksal HW, Lum J;  
 DR WPI: 91-117528/16.  
 DR N-PSDB; Q11289.  
 PT Human delta virus-specific peptide(s) - are used to detect and  
 PT treat human delta virus infections in humans  
 PS Disclosure; page 9; 18pp; English.  
 CC This polypeptide comprises residues 20-74 of the hepatitis delta virus  
 CC (HDV), a region corresp. to the delta antigenic epitope. A reagent  
 CC contg. it can be used to detect HDV infections by contacting with a  
 CC body sample and testing for an immunological reaction. It can also be  
 CC used to treat HDV infections when expressed as a high-yield fusion  
 CC protein in E.coli cells. An efficient mimic of the native protein is  
 CC produced. 55 AA;  
 SQ Sequence 55 AA;

Query Match 63.3%; Score 31; DB 1; Length 55;  
 Best Local Similarity 87.5%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
 |||||  
 DB 17 KTKKFLKK 24

RESULT 14  
 P82521 ID P82521 standard; Protein: 221 AA.  
 AC P82521;  
 DT 19-APR-1991 (first entry)  
 DE Hepatitis D Virus ORF5-encoded delta polypeptide.  
 KW Hepatitis D Virus; HDV; vaccine; ORF 5.  
 OS Hepatitis D Virus.  
 FH Key Location/Qualifiers  
 FT misc\_difference 104  
 FT /label= Arg, Lys  
 FT misc\_difference 160  
 FT /label= Asn, Ile  
 FT misc\_difference 179  
 FT /label= Leu, Gln  
 FT misc\_difference 199  
 FT /label= Gln, Arg  
 FT misc\_difference 203  
 FT /label= Trp, Amber stop  
 PN EP-251575-A.  
 PD 07-JAN-1988.  
 PF 17-JUN-1987; 305391.  
 PR 17-JUN-1986; US-875337.  
 PR 22-MAY-1987; US-053991.  
 PA (CHIR-) CHIRON CORP.  
 PI Houghton M, Wang KS, Choo QL, Weiner AJ, Overby LR;  
 DR WPI: 88-001440/01.  
 DR N-PSDB; N82172, N82173.  
 PT Nucleotide sequence of HDV (hepatitis D virus) - used for prodn. of  
 PT probes, polypeptides and antibodies for use in prodn. of HDV

PT diagnostics or vaccines.  
 PS Claim 1; Fig 3; 37pp; English.  
 CC There are several heterogeneities in the nucleotide sequence of ORF  
 CC 5, identified by comparing different clones containing ORF 5.  
 CC Some of the heterogeneities result in amino acid substitutions such  
 CC that ORF 5 encodes a family of closely related polypeptides. The  
 CC heterogeneity at nucleotide position 608 means that either an amber  
 CC stop codon or a Trp codon is specified at position 203 of the  
 CC polypeptide. This may be the difference between p24delta and  
 CC p27delta, the two viral polypeptides which are thought to be coded  
 CC for by ORF 5. The polypeptides are useful for diagnosing HDV  
 CC infections or for producing antibodies and vaccines.  
 CC See also N82174.  
 SQ Sequence 221 AA;

Query Match 63.3%; Score 31; DB 1; Length 221;  
 Best Local Similarity 87.5%; Pred. NO. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
 |||||  
 Db 43 KTKKFLKK 50

## RESULT 15

R64784  
 ID R64784 standard; peptide; 27 AA.  
 AC R64784;  
 DT 24-AUG-1995 (first entry)  
 DE Amphipathic peptide #15, for treating a pulmonary disease state.  
 KW Amphipathic peptide; pulmonary disease; resistant; bacteria;  
 KW proteolytic digestion; methylation; glyoxylation; cystic fibrosis;  
 KW CF; neoplasia; pneumonia; bronchitis; lytic activity; lysis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT modified\_site 1  
 FT /note= "Opt. methylated or glyoxylated"  
 FT modified\_site 1..27  
 FT /note= "Lys residues are opt. methylated and/or Arg  
 residues are glyoxylated"  
 PN W09428921-A.  
 PD 22-DEC-1994.  
 PF 02-JUN-1994; U06176.  
 PR 04-JUN-1993; US-039620.  
 PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.  
 PI Jaynes JM, Julian GR;  
 DR WPI; 95-036106/05.  
 PT Treatment of pulmonary disease states such as cystic fibrosis  
 PT by admin. of a non-naturally occurring amphipathic peptide  
 PS Claim 21; Page 48; 54pp; English.  
 CC The sequences given in R64770-806 and R76077 represent non-naturally  
 CC occurring amphipathic peptides which may be used for treating  
 CC pulmonary disease states. These peptides have enhanced resistance  
 CC to proteolytic digestion due to methylation of the epsilon-amino  
 CC group of Lys residues or glyoxylation of the side chains of Arg  
 CC residues, and methylation or glyoxylation of the N-terminal amino  
 CC acid. The peptides are prepared by standard methods of solid phase  
 CC synthesis and may be used in the treatment of cystic fibrosis (CF),  
 CC neoplasia, pneumonia, bronchitis, etc. The peptides pref. have a  
 CC lytic activity, thereby lysing pathogenic bacteria, virally infected  
 CC cells and transformed cells as well as treating the epithelial cell  
 CC defect of CF.  
 SQ Sequence 27 AA;

Query Match 63.3%; Score 31; DB 1; Length 27;  
 Best Local Similarity 75.0%; Pred. NO. 17;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
 |||||  
 Db 1 KKKKFKKK 8

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:50 ; Search time 82.37 Seconds  
(without alignments)  
1.198 Million cell updates/sec

Title: US-09-124-280A-36  
Perfect score: 49  
Sequence: 1 IKTKKFLKKT 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	10	1	US-08-049-871-7
2	49	100.0	10	1	US-07-819-893-7
3	49	100.0	10	1	US-08-280-397-7
4	49	100.0	10	2	US-08-218-026-48
5	49	100.0	10	2	US-08-653-632-48
6	49	100.0	10	2	US-08-456-1128-36
7	49	100.0	10	3	PCT-US94-01234-46
8	33.5	68.4	11	1	US-08-049-871-6
9	33.5	68.4	11	1	US-07-819-893-6
10	33.5	68.4	11	1	US-08-280-397-6
11	33.5	68.4	11	2	US-08-456-1128-35
12	31	63.3	27	1	US-08-231-730A-15
13	31	63.3	32	1	US-08-231-730A-16
14	31	63.3	37	1	US-08-231-730A-17
15	31	63.3	27	1	US-08-231-730A-21
16	31	63.3	32	1	US-08-231-730A-22
17	31	63.3	37	1	US-08-231-730A-23
18	31	63.3	968	1	US-08-434-730-14
19	31	63.3	27	1	US-08-427-001C-15
20	31	63.3	32	1	US-08-427-001C-16
21	31	63.3	37	1	US-08-427-001C-17
22	31	63.3	27	1	US-08-427-001C-21
23	31	63.3	32	1	US-08-427-001C-22
24	31	63.3	37	1	US-08-427-001C-23
25	31	63.3	27	1	US-08-457-798-15
26	31	63.3	32	1	US-08-457-798-16
27	31	63.3	37	1	US-08-457-798-17
28	31	63.3	27	1	US-08-457-798-21
29	31	63.3	32	1	US-08-457-798-22
30	31	63.3	37	1	US-08-457-798-23
31	31	63.3	27	2	US-08-457-171-15
32	31	63.3	32	2	US-08-457-171-16
33	31	63.3	37	2	US-08-457-171-17
34	31	63.3	27	2	US-08-457-171-21
35	31	63.3	32	2	US-08-457-171-22
36	31	63.3	37	2	US-08-457-171-23
37	31	63.3	27	3	PCT-US94-06176-15
38	31	63.3	32	3	PCT-US94-06176-16
39	31	63.3	37	3	PCT-US94-06176-17

40 31 63.3 27 3 PCT-US94-06176-21 Sequence 21, Appl  
41 31 63.3 32 3 PCT-US94-06176-22 Sequence 22, Appl  
42 31 63.3 37 3 PCT-US94-06176-23 Sequence 23, Appl  
43 31 63.3 27 3 PCT-US94-12550-15 Sequence 15, Appl  
44 31 63.3 32 3 PCT-US94-12550-16 Sequence 16, Appl  
45 31 63.3 37 3 PCT-US95-09339-23 Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-08-049-871-7  
; Sequence 7, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-049-871-7

Query Match 100.0%; Score 49; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.015; 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0;

QY 1 IKTKKFLKKT 10  
Db 1 IKTKKFLKKT 10

RESULT 2  
US-07-819-893-7  
; Sequence 7, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
US-07-819-893-7

Query Match 100.0%; Score 49; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKKT 10  
DB 1 IKTKKFLKKT 10

RESULT 3  
US-08-280-397-7  
;; Sequence 7, Application US/08280397  
;; Patent No. 5589459  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
US-08-280-397-7

Query Match 100.0%; Score 49; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKKT 10  
DB 1 IKTKKFLKKT 10

RESULT 4  
US-08-218-026-48  
;; Sequence 48, Application US/08218026  
;; Patent No. 5786324  
;; GENERAL INFORMATION:  
;; APPLICANT: Gray, Beulah  
;; APPLICANT: Haseman, Judith R.  
;; TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
;; TITLE OF INVENTION: Activity and Endotoxin Neutralizing Activity for Gram  
;; TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
;; NUMBER OF SEQUENCES: 60  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Merchant & Gould  
;; STREET: 3100 No. 5786324west Center  
;; CITY: Minneapolis  
;; STATE: MN  
;; COUNTRY: USA  
;; ZIP: 55402

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/218,026  
;; FILING DATE: 24-MAR-1994  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kowalchuk, Katherine M.  
;; REGISTRATION NUMBER: 36,848  
;; REFERENCE/DOCKET NUMBER: 600.286US01  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 612-332-5300  
;; TELEFAX: 612-332-9081  
;; INFORMATION FOR SEQ ID NO: 48:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-218-026-48

Query Match 100.0%; Score 49; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKFLKKT 10  
|||||  
Db 1 IKTKFLKKT 10

## RESULT 5

US-08-653-632-48  
; Sequence 48, Application US/08653632  
; Patent No. 5830860  
; GENERAL INFORMATION:  
; APPLICANT: GRAY, Beulah  
; APPLICANT: HASEMAN, Judith R.  
; APPLICANT: MAYO, Kevin  
; TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
; STREET: 3100 No. 5830860 West Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/653,632  
; FILING DATE: 24-MAY-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/218026  
; FILING DATE: 24-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kettelberger, Denise M  
; REGISTRATION NUMBER: 33,924  
; REFERENCE/DOCKET NUMBER: 600,286US11  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/371-5268  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: Internal  
; ORIGINAL SOURCE:  
US-08-653-632-48

Query Match 100.0%; Score 49; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKFLKKT 10  
|||||  
Db 1 IKTKFLKKT 10

## RESULT 6

US-08-456-112B-36  
; Sequence 36, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-36

Query Match 100.0%; Score 49; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKFLKKT 10  
|||||  
Db 1 IKTKFLKKT 10

## RESULT 7

PCT-US94-01234-46  
; Sequence 46, Application PC/TUS9401234  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE  
; TITLE OF INVENTION: BINDING SITES  
; NUMBER OF SEQUENCES: 76  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01234  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/084,542  
; FILING DATE: 28-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/012,566  
; FILING DATE: 02-FEB-1993  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal

PCT-US94-01234-46

Query Match 100.0%; Score 49; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.015; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKKT 10  
|||||  
Db 1 IKTKKFLKKT 10

RESULT 8

US-08-049-871-6  
; Sequence 6, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:

CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:

ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

TOPOLOGY: circular

US-08-049-871-6

Query Match 68.4%; Score 33.5; DB 1; Length 11;  
Best Local Similarity 90.0%; Pred. No. 4.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKTK-KFLKK 9  
|||||  
Db 1 IKTKKFLKK 10

RESULT 9

US-07-819-893-6  
; Sequence 6, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification

; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/819,893

FILING DATE: 19920115

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: AMINO ACID

TOPOLOGY: circular

US-07-819-893-6

Query Match 68.4%; Score 33.5; DB 1; Length 11;  
Best Local Similarity 90.0%; Pred. No. 4.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKTK-KFLKK 9

|||||

Db 1 IKTKKFLKK 10

RESULT 10

US-08-280-397-6

; Sequence 6, Application US/08280397

; Patent No. 5589459

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: Synthetic Peptides for Detoxification

; TITLE OF INVENTION: of Bacterial Endotoxins and for the

; TITLE OF INVENTION: Prevention and Treatment of Septic

; TITLE OF INVENTION: Shock

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson &amp; Costigan, P.C.

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/280,397

FILING DATE: 07/26/94



CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-6

Query Match 68.4%; Score 33.5; DB 1; Length 11;  
Best Local Similarity 90.0%; Pred. No. 4.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKTK-KFLKK 9  
|||||  
DB 1 IKTKCKFLKK 10

RESULT 11  
US-08-456-112B-35  
Sequence 35, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-35

Query Match 68.4%; Score 33.5; DB 2; Length 11;  
Best Local Similarity 90.0%; Pred. No. 4.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKTK-KFLKK 9  
|||||  
DB 1 IKTKCKFLKK 10

RESULT 12  
US-08-231-730A-15  
Sequence 15, Application US/08231730A  
Patent No. 5561107  
GENERAL INFORMATION:  
APPLICANT: JAYNES, JESSE M.  
APPLICANT: JULIAN, GORDON R.  
TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STEVEN J. HULTQUIST  
ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW  
STREET: 200 PARK DRIVE, SUITE 210  
STREET: P.O. BOX 14329  
CITY: RESEARCH TRIANGLE PARK  
STATE: NORTH CAROLINA  
COUNTRY: USA  
ZIP: 27709

COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE  
COMPUTER: APPLE MACINTOSH  
OPERATING SYSTEM: MACINTOSH  
SOFTWARE: M.S. WORD 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,730A  
FILING DATE: 04-20-94  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/225,476  
FILING DATE: 04-08-94  
APPLICATION NUMBER: 08/039,620  
FILING DATE: 06-04-93  
APPLICATION NUMBER: 08/148,491  
FILING DATE: 11-08-93  
APPLICATION NUMBER: 08/148,889  
FILING DATE: 11-08-93  
ATTORNEY/AGENT INFORMATION:  
NAME: HULTQUIST, STEVEN J.  
REGISTRATION NUMBER: 28021  
REFERENCE/DOCKET NUMBER: 4013-106  
TELEPHONE: (919)990-9531  
TELEFAX: (919)990-9532  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE:  
DESCRIPTION: PEPTIDE  
HYPOTHETICAL: NO  
FRAGMENT TYPE: COMPLETE PEPTIDE  
ORIGINAL SOURCE: SYNTHETIC  
IMMEDIATE SOURCE: SYNTHETIC  
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED  
US-08-231-730A-15

Query Match 63.3%; Score 31; DB 1; Length 27;  
Best Local Similarity 75.0%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
|||||  
DB 1 KKKKFLKK 8

RESULT 13

US-08-231-730A-16  
; Sequence 16, Application US/08231730A  
; Patent No. 5561107  
; GENERAL INFORMATION:  
; APPLICANT: JAYNES, JESSE M.  
; APPLICANT: JULIAN, GORDON R.  
; TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST AN  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STEVEN J. HULTQUIST  
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW  
; STREET: 200 PARK DRIVE, SUITE 210  
; STREET: P.O. BOX 14329  
; CITY: RESEARCH TRIANGLE PARK  
; STATE: NORTH CAROLINA  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE  
; COMPUTER: APPLE MACINTOSH  
; OPERATING SYSTEM: MACINTOSH  
; SOFTWARE: M.S. WORD 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,730A  
; FILING DATE: 04-20-94  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/225,476  
; FILING DATE: 04-08-94  
; APPLICATION NUMBER: 08/039,620  
; FILING DATE: 06-04-93  
; APPLICATION NUMBER: 08/148,491  
; FILING DATE: 11-08-93  
; APPLICATION NUMBER: 08/148,889  
; FILING DATE: 11-08-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HULTQUIST, STEVEN J.  
; REGISTRATION NUMBER: 28021  
; REFERENCE/DOCKET NUMBER: 4013-106  
; TELEPHONE: (919)990-9531  
; TELEFAX: (919)990-9532  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PEPTIDE  
; DESCRIPTION: NO  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: COMPLETE PEPTIDE  
; ORIGINAL SOURCE: SYNTHETIC  
; IMMEDIATE SOURCE: SYNTHETIC  
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED  
US-08-231-730A-16

Query Match 63.3%; Score 31; DB 1; Length 32;  
Best Local Similarity 75.0%; Pred. No. 26;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
; | |||:|  
Db 1 KKKKEVKK 8

RESULT 14  
US-08-231-730A-17  
; Sequence 17, Application US/08231730A  
; Patent No. 5561107  
; GENERAL INFORMATION:  
; APPLICANT: JAYNES, JESSE M.  
; APPLICANT: JULIAN, GORDON R.  
; TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST AN

; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STEVEN J. HULTQUIST  
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW  
; STREET: 200 PARK DRIVE, SUITE 210  
; STREET: P.O. BOX 14329  
; CITY: RESEARCH TRIANGLE PARK  
; STATE: NORTH CAROLINA  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE  
; COMPUTER: APPLE MACINTOSH  
; OPERATING SYSTEM: MACINTOSH  
; SOFTWARE: M.S. WORD 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,730A  
; FILING DATE: 04-20-94  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/225,476  
; FILING DATE: 04-08-94  
; APPLICATION NUMBER: 08/039,620  
; FILING DATE: 06-04-93  
; APPLICATION NUMBER: 08/148,491  
; FILING DATE: 11-08-93  
; APPLICATION NUMBER: 08/148,889  
; FILING DATE: 11-08-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HULTQUIST, STEVEN J.  
; REGISTRATION NUMBER: 28021  
; REFERENCE/DOCKET NUMBER: 4013-106  
; TELEPHONE: (919)990-9531  
; TELEFAX: (919)990-9532  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PEPTIDE  
; DESCRIPTION: NO  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: COMPLETE PEPTIDE  
; ORIGINAL SOURCE: SYNTHETIC  
; IMMEDIATE SOURCE: SYNTHETIC  
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED  
US-08-231-730A-17

Query Match 63.3%; Score 31; DB 1; Length 37;  
Best Local Similarity 75.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
; | |||:|  
Db 1 KKKKEVKK 8

RESULT 15  
US-08-231-730A-21  
; Sequence 21, Application US/08231730A  
; Patent No. 5561107  
; GENERAL INFORMATION:  
; APPLICANT: JAYNES, JESSE M.  
; APPLICANT: JULIAN, GORDON R.  
; TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STEVEN J. HULTQUIST  
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW  
; STREET: 200 PARK DRIVE, SUITE 210  
; STREET: P.O. BOX 14329  
; CITY: RESEARCH TRIANGLE PARK

STATE: NORTH CAROLINA  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE  
COMPUTER: APPLE MACINTOSH  
OPERATING SYSTEM: MACINTOSH  
SOFTWARE: M.S. WORD 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,730A  
FILING DATE: 04-20-94  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/225,476  
FILING DATE: 04-08-94  
APPLICATION NUMBER: 08/039,620  
FILING DATE: 06-04-93  
APPLICATION NUMBER: 08/148,491  
FILING DATE: 11-08-93  
APPLICATION NUMBER: 08/148,889  
FILING DATE: 11-08-93  
ATTORNEY/AGENT INFORMATION:  
NAME: HULTQUIST, STEVEN J.  
REGISTRATION NUMBER: 28021  
REFERENCE/DOCKET NUMBER: 4013-106  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)990-9531  
TELEFAX: (919)990-9532  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE:  
DESCRIPTION: PEPTIDE  
HYPOTHETICAL: NO  
FRAGMENT TYPE: COMPLETE PEPTIDE  
ORIGINAL SOURCE: SYNTHETIC  
IMMEDIATE SOURCE: SYNTHETIC  
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED  
US-08-231-730A-21

Query Match 63.3%; Score 31; DB 1; Length 27;  
Best Local Similarity 75.0%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 KKKKELKK 9  
Db 1 KKKKFKK 8

Search completed: September 7, 1999, 23:07:50  
Job time: 1748 sec

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	37	75.5	1000	2	JE0110	Dis3p protein - hu	
2	36	73.5	239	2	A46603	decidual prolactin	
3	35	71.4	1161	2	B70172	DNA polymerase III	
4	35	71.4	840	2	A27832	cell division control	
5	34	69.4	569	2	A45824	trophozoite cysteine	
6	34	69.4	1018	2	JC4375	DNA-directed DNA p	
7	34	69.4	372	2	F64310	hypothetical protei	
8	34	69.4	1193	2	S68218	botulinum neurotox	
9	34	69.4	782	2	S04047	finger protein zfy	
10	34	69.4	783	2	A31491	sex-determining re	
11	34	69.4	1980	2	S54307	myosin heavy chain	
12	33	67.3	323	2	S39743	phosphotransacetyl	
13	33	67.3	132	2	F64487	hypothetical prote	
14	33	67.3	249	2	D64494	conserved hypothet	
15	33	67.3	202	2	B71509	probable phosphata	
16	33	67.3	287	2	B64201	hypothetical prote	
17	33	67.3	58	2	H64301	hypothetical prote	
18	33	67.3	232	2	A64366	hypothetical prote	
19	33	67.3	517	2	S49795	alpha-1,2-mannosyl	
20	33	67.3	312	2	S50430	hypothetical prote	
21	33	67.3	404	2	S34031	KTR3 protein - yea	
22	33	67.3	404	2	S63352	probable membrane	
23	33	67.3	568	2	A55377	CPE-binding protei	
24	33	67.3	667	2	A40713	cyclicin I - bovine	
25	33	67.3	226	2	A70436	hypothetical prote	
26	33	67.3	219	2	D71001	hypothetical prote	
27	32	65.3	1024	1	RN208F	DNA-directed RNA p	
28	32	65.3	741	2	B54308	phospholipase A2 c	
29	32	65.3	806	2	E64221	phenylalanine--trn	
30	32	65.3	122	2	S58649	ribosomal protein	
31	32	65.3	121	2	S61635	ribosomal protein	
32	32	65.3	125	2	JU0179	heparin-binding pr	
33	32	65.3	130	2	A30033	development-specif	
34	32	65.3	128	2	JC2119	heparin-binding pr	
35	32	65.3	128	2	JC2121	heparin-binding pr	
36	32	65.3	128	2	S52084	ribosomal protein	
37	32	65.3	128	2	JC2120	heparin-binding pr	
38	32	65.3	657	2	A29454	kno3-associated h1	
39	32	65.3	2376	2	S48405	probable membrane	

A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: B70172  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1161 <LUE>  
 A:Cross-references: GB:AF001159; GB:AF000783; NID:g2688496; PID:g2688500; TIGR:BB0579  
 A:Experimental source: strain B31  
 C:Superfamily: DNA-directed DNA polymerase III alpha chain

Query Match 71.4%; Score 35; DB 2; Length 1161;  
 Best Local Similarity 87.5%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
 || |||||  
 Db 89 KTSKFLKK 96

RESULT 4  
 A27832  
 cell division control protein CDC16 - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: protein YKL022c  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 19-May-1989 #sequence\_revision 19-May-1989 #text\_change 09-Apr-1998  
 C:Accession: A27832; S37839  
 R:Ichio, T.; Wickner, R.B.  
 Nucleic Acids Res. 15, 8439-8450, 1987  
 A:Title: Metal-binding, nucleic acid-binding finger sequences in the CDC16 gene of *Sacch*  
 A:Reference number: A27832; MUID:88040465  
 A:Accession: A27832  
 A:Molecule type: DNA  
 A:Residues: 1-840 <ICH>  
 A:Cross-references: EMBL:X06165; NID:g3477; PID:g3478  
 R:Rieger, M.  
 submitted to the Protein Sequence Database, March 1994  
 A:Reference number: S37832  
 A:Accession: S37839  
 A:Molecule type: DNA  
 A:Residues: 1-840 <RIE>  
 A:Cross-references: EMBL:Z28022; NID:g486014; PID:g486015; MIPS:YKL022c  
 A:Experimental source: strain S286C  
 C:Genetics:  
 A:Gene: SGD:CDC16  
 A:Cross-references: SGD:S0001505; MIPS:YKL022c  
 A:Map position: III  
 C:Superfamily: tetratricopeptide repeat homology  
 C:Keywords: cell cycle control; nucleus  
 F:456-489/Domain: tetratricopeptide repeat homology <TT1>  
 F:457-530/Domain: tetratricopeptide repeat homology <TT2>  
 F:531-564/Domain: tetratricopeptide repeat homology <TT3>  
 F:565-598/Domain: tetratricopeptide repeat homology <TT4>  
 F:599-632/Domain: tetratricopeptide repeat homology <TT5>  
 F:633-666/Domain: tetratricopeptide repeat homology <TT6>

Query Match 71.4%; Score 35; DB 2; Length 840;  
 Best Local Similarity 66.7%; Pred. No. 51;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKKFLKK 9  
 : |||||  
 Db 650 VKAKKFLKK 658

RESULT 5  
 A45624  
 trophozoite cysteine proteinase - *Plasmodium falciparum*  
 C:Species: *Plasmodium falciparum*  
 C:Date: 22-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Sep-1998  
 C:Accession: A45624

R:Rosenthal, P.J.; Nelson, R.G.  
 Mol. Biochem. Parasitol. 51, 143-152, 1992  
 A:Title: Isolation and characterization of a cysteine proteinase gene of *Plasmodium f*  
 A:Reference number: A45624; MUID:92228005  
 A:Accession: A45624  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-569 <ROS>  
 A:Cross-references: GB:M81341; GB:M80590; NID:g160247; PID:g160248  
 A:Note: sequence extracted from NCBI backbone (NCBI:98870, NCBI:98874)  
 C:Superfamily: trophozoite cysteine proteinase

Query Match 69.4%; Score 34; DB 2; Length 569;  
 Best Local Similarity 87.5%; Pred. No. 54;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
 || |||||  
 Db 23 KKKKFLKK 30

RESULT 6  
 JC4375  
 DNA-directed DNA polymerase (EC 2.7.7.7) gamma, mitochondrial - fission yeast (*Schizo*  
 N:Alternate names: DNA polymerase gamma  
 C:Species: *Schizosaccharomyces pombe*  
 C:Date: 10-Jan-1996 #sequence\_revision 08-Feb-1996 #text\_change 17-Mar-1999  
 C:Accession: JC4375; S57641  
 R:Ropp, P.A.; Copeland, W.C.  
 Gene 165, 103-107, 1995  
 A:Title: Characterization of a new DNA polymerase from *Schizosaccharomyces pombe*: A p  
 A:Reference number: JC4375; MUID:96084961  
 A:Accession: JC4375  
 A:Molecule type: DNA  
 A:Residues: 1-1018 <ROP>  
 A:Cross-references: EMBL:247976  
 C:Comment: This enzyme is unique among the eukaryotic DNA polymerase, and is the only  
 C:Genetics:  
 A:Gene: pol gamma  
 A:Map position: III  
 C:Superfamily: DNA-directed DNA polymerase gamma chain  
 C:Keywords: DNA replication; mitochondrion; nucleotidyltransferase  
 F:169-179/Region: exonuclease pattern A  
 F:223-230/Region: exonuclease pattern B  
 F:351-355/Region: exonuclease pattern C

Query Match 69.4%; Score 34; DB 2; Length 1018;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFL 7  
 |||||  
 Db 20 IKTKKFL 26

RESULT 7  
 F64310  
 hypothetical protein MJ0086 - *Methanococcus jannaschii*  
 C:Species: *Methanococcus jannaschii*  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Sep-1998  
 C:Accession: F64310  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak  
 ; Reich, C.L.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,  
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
 A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannas*  
 A:Reference number: A64300; MUID:96337999  
 A:Accession: F64310  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA

A:Residues: 1-372 <BUL>  
A:Cross-references: GB:U67466; GB:L77117; NID:g1590867; PID:g1498850; TIGR:MJ0086  
C:Genetics:  
A:Map position: FOR80788-81906  
C:Superfamily: hypothetical protein MJ0086

Query Match 69.4%; Score 34; DB 2; Length 372;  
Best Local Similarity 75.0%; Pred. No. 36;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
:||||:|  
Db 209 ETKKFKK 216

RESULT 8  
S68218  
botulinum neurotoxin type A, nontoxic/nonhemagglutinating - Clostridium botulinum (strain  
C:Species: Clostridium botulinum  
A:Variety: strain NIH  
C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 12-Jun-1998  
C:Accession: S68218; S74301  
R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.  
FEBS Lett. 376, 41-44, 1995  
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin components  
A:Reference number: S67988; MUID:96096783  
A:Accession: S68218  
A:Molecule type: DNA  
A:Residues: 1-1193 <FUJ>  
A:Cross-references: EMBL:D67030; NID:g2160224; PID:d1011710; PID:g1132475  
A:Experimental source: strain NIH  
A:Accession: S74301  
A:Molecule type: protein  
A:Residues: 1-13,145-155 <FUJ>  
A:Experimental source: strain NIH  
C:Genetics:  
A:Gene: ant  
C:Keywords: neurotoxin  
F:1-1193/Product: botulinum neurotoxin type A, nontoxic/nonhemagglutinating #status expected

Query Match 69.4%; Score 34; DB 2; Length 1193;  
Best Local Similarity 66.7%; Pred. No. 1,1e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKK 9  
||||:|  
Db 778 IKTKFIQK 786

RESULT 9  
S04047  
finger protein zfy-1 - mouse  
N:Alternate names: probable testis-determining factor  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 10-Sep-1997  
C:Accession: S04047  
R:Ashworth, A.; Swift, S.; Affara, N.  
Nucleic Acids Res. 17, 2864, 1989  
A:Title: Sequence of cDNA for murine zfy-1, a candidate for Tdy.  
A:Reference number: S04047; MUID:89240049  
A:Accession: S04047  
A:Molecule type: mRNA  
A:Residues: 1-782 <ASH>  
A:Cross-references: EMBL:X14382; NID:g55478; PID:g55479  
C:Genetics:  
A:Gene: zfy-1  
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 69.4%; Score 34; DB 2; Length 782;  
Best Local Similarity 75.0%; Pred. No. 73;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KTKKFLKK 9  
||||:|  
Db 413 KTKRFLKR 420

RESULT 10  
A31491  
sex-determining region protein zfy-2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jul-1989 #sequence\_revision 13-Jul-1989 #text\_change 20-Mar-1998  
C:Accession: A31491  
R:Mardon, G.; Page, D.C.  
Cell 56, 765-770, 1989  
A:Title: The sex-determining region of the mouse Y chromosome encodes a protein with  
A:Reference number: A31491; MUID:89168416  
A:Accession: A31491  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-783 <MAR>  
A:Cross-references: GB:M24401; NID:g598331; PID:g598332  
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 69.4%; Score 34; DB 2; Length 783;  
Best Local Similarity 75.0%; Pred. No. 73;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
||||:|  
Db 413 KTKRFLKR 420

RESULT 11  
S54307  
myosin heavy chain - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
C:Accession: S54307  
R:Reinhardt, J.; Scheel, A.A.; Diekmann, D.; Hall, A.; Ruppert, C.; Baehler, M.  
EMBO J. 14, 697-704, 1995  
A:Title: A novel type of myosin implicated in signalling by rho family GTPases.  
A:Reference number: S54307; MUID:95188874  
A:Accession: S54307  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1980 <REI>  
A:Cross-references: EMBL:X77609; NID:g639998; PID:g639999  
C:Superfamily: myosin motor domain homology  
C:Keywords: P-loop  
F:149-942/Domain: myosin motor domain homology <MMOT>  
F:239-246/Region: nucleotide-binding motif A (P-loop)

Query Match 69.4%; Score 34; DB 2; Length 1980;  
Best Local Similarity 77.8%; Pred. No. 1.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKKT 10  
||||:|  
Db 936 KTKVFLKET 944

RESULT 12  
S39743  
phosphotransacetylase pta - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 07-Oct-1994 #sequence\_revision 22-Nov-1996 #text\_change 17-Mar-1999  
C:Accession: S39743; D69683  
R:Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu, A.; Rapoport, G.; Danchin, A.  
Mol. Microbiol. 10, 371-384, 1993

A:Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fd  
A:Reference number: S39655; MUID:95020537  
A:Accession: S39743  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-323 <GLA>  
A:Cross-references: EMBL:X73124; NID:g413923; PID:g580883  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993  
R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berton  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1998  
C:Accession: D64494  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999  
A:Accession: D64494  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-323 <KUN>  
A:Cross-references: GB:299123; GB:AL009126; NID:g2636240; PID:e1186266; PID:g2636302  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: pta  
A:Start codon: GTG  
C:Superfamily: phosphate acetyltransferase

Query Match 67.3%; Score 33; DB 2; Length 323;  
Best Local Similarity 70.0%; Pred. No. 49;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKFLKKT 10  
||||:||||  
DB 140 IKTKGVKKT 149

RESULT 13  
F64487  
hypothetical protein MJ1503 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1998  
C:Accession: F64487  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
rson, J.D.; Sadov, P.W.; Hann, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999  
A:Accession: F64487  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-132 <BUL>  
A:Cross-references: GB:U67591; GB:L77117; NID:g1592136; PID:g1500391; TIGR:MJ1503  
C:Genetics:  
A:Map position: REV1474984-1474586  
C:Superfamily: conserved hypothetical protein MJ1082

Query Match 67.3%; Score 33; DB 2; Length 132;  
Best Local Similarity 66.7%; Pred. No. 21;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKKT 10

Db 94 ETKKIIRKT 102  
|||:||||

## RESULT 14

D64494  
conserved hypothetical protein MJ1557 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1998  
C:Accession: D64494  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
rson, J.D.; Sadov, P.W.; Hann, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999  
A:Accession: D64494  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-249 <BUL>  
A:Cross-references: GB:U67596; GB:L77117; NID:g1592181; PID:g1500450; TIGR:MJ1557  
C:Genetics:  
A:Map position: FOR1533646-1534395  
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1557

Query Match 67.3%; Score 33; DB 2; Length 249;  
Best Local Similarity 77.8%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KTKKFLKKT 10  
|||:||||

Db 174 KTHKFLDKT 182

## RESULT 15

B71509  
probable phosphatase/kinase - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 21-Nov-1998  
C:Accession: B71509  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t  
A:Reference number: A71570; MUID:99000809  
A:Accession: B71509  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-202 <ARN>  
A:Cross-references: GB:AE001322; GB:AE001273; NID:g3328916; PID:g3328928  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: yacE

Query Match 67.3%; Score 33; DB 2; Length 202;  
Best Local Similarity 60.0%; Pred. No. 31;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKFLKKT 10  
|||:||||

Db 141 IKTKFLKKT 150

Search completed: September 7, 1999, 23:22:52  
Job time: 959 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:49 ; Search time 82.37 Seconds  
(without alignments)  
1.078 Million cell updates/sec

Title: US-09-124-280A-34  
Perfect score: 53  
Sequence: 1 CKKLFCKCT 9

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description				
1	53	100.0	10	1	US-08-049-871-4	Sequence 4, Appl			
2	53	100.0	9	1	US-08-049-871-5	Sequence 5, Appl			
3	53	100.0	10	1	US-07-819-893-4	Sequence 4, Appl			
4	53	100.0	9	1	US-07-819-893-5	Sequence 5, Appl			
5	53	100.0	10	1	US-08-280-397-4	Sequence 4, Appl			
6	53	100.0	9	1	US-08-280-397-5	Sequence 5, Appl			
7	53	100.0	10	2	US-08-456-112B-33	Sequence 33, Appl			
8	53	100.0	9	2	US-08-456-112B-34	Sequence 34, Appl			
9	32	60.4	1141	1	US-08-363-300-2	Sequence 2, Appl			
10	32	60.4	1706	2	US-08-459-568-2	Sequence 2, Appl			
11	32	60.4	86	2	US-08-459-568-82	Sequence 82, Appl			
12	32	60.4	1706	2	US-08-399-411-2	Sequence 2, Appl			
13	32	60.4	86	2	US-08-399-411-82	Sequence 82, Appl			
14	31.5	59.4	10	1	US-08-097-830E-18	Sequence 18, Appl			
15	31.5	59.4	10	2	US-08-456-112B-18	Sequence 18, Appl			
16	30	56.6	32	1	US-08-361-920-10	Sequence 10, Appl			
17	30	56.6	514	1	US-08-361-920-21	Sequence 21, Appl			
18	30	56.6	740	1	US-08-276-099A-12	Sequence 12, Appl			
19	30	56.6	32	1	US-08-479-939-10	Sequence 10, Appl			
20	30	56.6	514	1	US-08-479-939-21	Sequence 21, Appl			
21	30	56.6	740	1	US-08-781-890-12	Sequence 12, Appl			
22	30	56.6	750	1	US-08-369-796-4	Sequence 4, Appl			
23	30	56.6	712	1	US-08-369-796-6	Sequence 6, Appl			
24	30	56.6	749	1	US-08-369-796-8	Sequence 8, Appl			
25	30	56.6	32	2	US-08-483-432-10	Sequence 10, Appl			
26	30	56.6	514	2	US-08-483-432-21	Sequence 21, Appl			
27	30	56.6	1719	2	US-08-459-568-4	Sequence 4, Appl			
28	30	56.6	1719	2	US-08-399-411-4	Sequence 4, Appl			
29	30	56.6	750	2	US-08-852-091-4	Sequence 4, Appl			
30	30	56.6	712	2	US-08-852-091-6	Sequence 6, Appl			
31	30	56.6	749	2	US-08-852-091-8	Sequence 8, Appl			
32	30	56.6	750	3	PCT-US95-17025-4	Sequence 4, Appl			
33	30	56.6	712	3	PCT-US95-17025-6	Sequence 6, Appl			
34	30	56.6	749	3	PCT-US95-17025-8	Sequence 8, Appl			
35	30	56.6	200	4	5188960-8	Patent No. 5188960			
36	29.5	55.7	374	1	US-08-450-393A-2	Sequence 2, Appl			
37	29.5	55.7	360	1	US-08-450-393A-4	Sequence 4, Appl			
38	29.5	55.7	347	2	US-08-461-244-3	Sequence 3, Appl			
39	29.5	55.7	374	3	PCT-US95-00476-2	Sequence 2, Appl			

ALIGNMENTS

RESULT 1

US-08-049-871-4

; Sequence 4, Application US/08049871

; Patent No. 5358933

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: Synthetic Peptides for Detoxification

; TITLE OF INVENTION: of Bacterial Endotoxins and for the

; TITLE OF INVENTION: Prevention and Treatment of Septic

; TITLE OF INVENTION: Shock

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson, Costigan & Hoare

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: DOS

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/049,871

; FILING DATE:

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/658,744

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Costigan, James V.

; REGISTRATION NUMBER: 25,669

; REFERENCE/DOCKET NUMBER: 576-001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 302-8989

; TELEFAX: (212) 302-8998

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10 amino acids

; TYPE: amino acid

; TOPOLOGY: circular

US-08-049-871-4

Query Match 100.0%; Score 53; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CKKLFCKCT 9

Db 1 CKKLFCKCT 9

RESULT 2

US-08-049-871-5

; Sequence 5, Application US/08049871

; Patent No. 5358933

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: Synthetic Peptides for Detoxification

; TITLE OF INVENTION: of Bacterial Endotoxins and for the

```

; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,871
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/658,744
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; US-08-049-871-5

```

```

Query Match 100.0%; Score 53; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CKKLFCKCT 9
    |||||
Db 1 CKKLFCKCT 9

```

```

RESULT 3
US-07-819-893-4
; Sequence 4, Application US/07819893
; Patent No. 5371186
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/819,893
; FILING DATE: 19920115
; CLASSIFICATION: 514

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: circular
; US-07-819-893-4

```

```

Query Match 100.0%; Score 53; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CKKLFCKCT 9
    |||||
Db 1 CKKLFCKCT 9

```

```

RESULT 4
US-07-819-893-5
; Sequence 5, Application US/07819893
; Patent No. 5371186
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/819,893
; FILING DATE: 19920115
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: circular
; US-07-819-893-5

```

```

Query Match 100.0%; Score 53; DB 1; Length 9;

```

Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKT 9  
| | | | | | | |  
Db 1 CKKLFCKT 9

RESULT 5  
US-08-280-397-4  
; Sequence 4, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: POPIO, Masslmo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-280-397-4

Query Match 100.0%; Score 53; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKT 9  
| | | | | | | |  
Db 1 CKKLFCKT 9

RESULT 6  
US-08-280-397-5  
; Sequence 5, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: POPIO, Masslmo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,397  
FILING DATE: 07/26/94  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-5

Query Match 100.0%; Score 53; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKKLFCKT 9  
| | | | | | | |  
Db 1 CKKLFCKT 9

RESULT 7  
US-08-456-112B-33  
; Sequence 33, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: POPIO, Masslmo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669

```

; REFERENCE/DOCKET NUMBER: 576-004
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-08-456-112B-33

```

Query Match 100.0%; Score 53; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 9; Conservative 0; Mismatches 0; Indels

Qy 1 CKKLFCKT 9  
| | | | | | | |  
Db 1 CKKLFCKT 9

```

RESULT      8
US-08-456-112B-34
: Sequence 34, Application US/08456112B
: Patent No. 5834430
:
: GENERAL INFORMATION:
: APPLICANT: Porro, Massimo
: TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hedman, Gibson & Costigan
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA

```

```
Query Match      100.0%; Score 53; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 9; Conservative 0; Mismatches 0; Indels
```

Qy 1 CKKLFCKT 9  
11111111

Dp 1 CKKLFCKT 9

## RESULT 9

```

US-08-363-300-2
: Sequence 2, Application US/08363300
: Patent No. 5700927
: GENERAL INFORMATION:
: APPLICANT: Zon, Leonard and Richardson, Paul
: TITLE OF INVENTION: Tbcd Gene and Uses Thereof
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30B
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/363,300
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Clark, Paul T.
: REGISTRATION NUMBER: 30,162
: REFERENCE/DOCKET NUMBER: 04590/009001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1141 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-363-300-2

```

Query Match 60.4%; Score 32; DB 1; Length 1141;  
Best Local Similarity 50.0%;  
Pred. No. 2.9e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels

QY	1	CKKLFCK	8
		:   :	
Db	54	CSSIFECK	61

RESULT 10  
US-08-459-568-2  
; Sequence 2, Application US/08459568  
; Patent No. 5811304  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; TITLE OF INVENTION: Zinc Finger Proteins  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,568  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 514

; PRIOR APPLICATION DATA: US 08/399,411  
 ; APPLICATION NUMBER: 06-MAR-1995  
 ; FILING DATE: 06-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Campbell, Cathryn A.  
 ; REGISTRATION NUMBER: 31,815  
 ; REFERENCE/DOCKET NUMBER: P-LJ 1264  
 ; TELEPHONE: (619) 535-9001  
 ; TELEFAX: (619) 535-8949  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1706 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-459-568-2

Query Match 60.4%; Score 32; DB 2; Length 1706;  
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKKT 9  
 Db 1156 CVQLFKVKT 1164

RESULT 11  
 US-08-459-568-82  
 ; Sequence 82, Application US/08459568  
 ; Patent No. 5811304  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Huang, Shi  
 ; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
 ; TITLE OF INVENTION: Zinc Finger Proteins  
 ; NUMBER OF SEQUENCES: 93  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Campbell and Flores  
 ; STREET: 4370 La Jolla Village Drive, Suite 700  
 ; CITY: San Diego  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 92122  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/459,568  
 ; FILING DATE: 02-JUN-1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/399,411  
 ; FILING DATE: 06-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Campbell, Cathryn A.  
 ; REGISTRATION NUMBER: 31,815  
 ; REFERENCE/DOCKET NUMBER: P-LJ 1264  
 ; TELEPHONE: (619) 535-9001  
 ; TELEFAX: (619) 535-8949  
 ; INFORMATION FOR SEQ ID NO: 82:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 86 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-459-568-82

Query Match 60.4%; Score 32; DB 2; Length 86;  
 Best Local Similarity 66.7%; Pred. No. 26;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CKKLFCKKT 9  
 Db 32 CVQLFKVKT 40

RESULT 12  
 US-08-399-411-2  
 ; Sequence 2, Application US/08399411  
 ; Patent No. 5831008  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Huang, Shi  
 ; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
 ; TITLE OF INVENTION: Zinc Finger Proteins  
 ; NUMBER OF SEQUENCES: 93  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Campbell and Flores  
 ; STREET: 4370 La Jolla Village Drive, Suite 700  
 ; CITY: San Diego  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 92122  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/399,411  
 ; FILING DATE: 06-MAR-1995  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Campbell, Cathryn A.  
 ; REGISTRATION NUMBER: 31,815  
 ; REFERENCE/DOCKET NUMBER: P-LJ 1264  
 ; TELEPHONE: (619) 535-9001  
 ; TELEFAX: (619) 535-8949  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1706 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-399-411-2

Query Match 60.4%; Score 32; DB 2; Length 1706;  
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKKLFCKKT 9  
 Db 1156 CVQLFKVKT 1164

RESULT 13  
 US-08-399-411-82  
 ; Sequence 82, Application US/08399411  
 ; Patent No. 5831008  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Huang, Shi  
 ; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
 ; TITLE OF INVENTION: Zinc Finger Proteins  
 ; NUMBER OF SEQUENCES: 93  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Campbell and Flores  
 ; STREET: 4370 La Jolla Village Drive, Suite 700  
 ; CITY: San Diego  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 92122

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/399,411  
; FILING DATE: 06-MAR-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1264  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 82:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 86 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-399-411-82

Query Match 60.4%; Score 32; DB 2; Length 86;  
Best Local Similarity 66.7%; Pred. No. 26;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKLKFCKT 9  
| : ||| |  
Db 32 CVOLEKVK 40

RESULT 14  
US-08-097-830E-18  
; Sequence 18, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular

; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and Cys  
; US-08-097-830E-18  
Query Match 59.4%; Score 31.5; DB 1; Length 10;  
Best Local Similarity 77.8%; Pred. No. 4.3;  
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CK-KLFCK 8  
| : ||| |  
Db 2 CKFKFKCK 10

RESULT 15  
US-08-456-112B-18  
; Sequence 18, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-18

Query Match 59.4%; Score 31.5; DB 2; Length 10;  
Best Local Similarity 77.8%; Pred. No. 4.3;  
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CK-KLFCK 8  
| : ||| |  
Db 2 CKFKFKCK 10

Search completed: September 7, 1999, 23:07:50  
Job time: 1748 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:15 ; Search time 71.09 Seconds  
(without alignments)  
3.976 Million cell updates/sec

Title: US-09-124-280A-36

Perfect score: 49

Sequence: 1 IKTKFLKKT 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	71.4	840	1 CC16_YEAST	P09798 saccharomyc
2	35	71.4	1161	1 DP3A_BORBU	O51526 borrellia bu
3	35	71.4	115	1 RL22_SCHPO	Q09668 schizosacch
4	34	69.4	569	1 CYSP_PLAFA	P25805 plasmodium
5	34	69.4	1018	1 DPOG_SCHPO	Q12704 schizosacch
6	34	69.4	372	1 Y086_METJA	Q57551 methanococc
7	34	69.4	782	1 ZFV1_MOUSE	P10925 mus musculu
8	34	69.4	783	1 ZFV2_MOUSE	P20862 mus musculu
9	33	67.3	667	1 CYL1_BOVIN	P35662 bos taurus
10	33	67.3	214	1 HS30_ONCTS	P42931 oncorhynch
11	33	67.3	404	1 KTR3_YEAST	P38130 saccharomyc
12	33	67.3	517	1 KTR7_YEAST	P40504 saccharomyc
13	33	67.3	322	1 PTA_BACSU	P39646 bacillus su
14	33	67.3	287	1 Y011_MYCGE	P47257 mycoplasma
15	33	67.3	58	1 Y016_METJA	Q60322 methanococc
16	33	67.3	232	1 Y529_METJA	Q57949 methanococc
17	33	67.3	312	1 YEC9_YEAST	P39988 saccharomyc
18	33	67.3	132	1 YF03_METJA	Q58998 methanococc
19	33	67.3	249	1 YF57_METJA	Q58998 methanococc
20	33	67.3	404	1 YNBB_YEAST	P53723 saccharomyc
21	32	65.3	824	1 DPOL_METVO	P52025 methanococc
22	32	65.3	293	1 ER25_HUMAN	Q15800 homo sapien
23	32	65.3	657	1 KNOB_PLAFN	P06719 plasmodium
24	32	65.3	741	1 PAZY_BRARE	P50392 brachydanio
25	32	65.3	120	1 R22A_YEAST	P05749 saccharomyc
26	32	65.3	121	1 R22B_YEAST	P56628 saccharomyc
27	32	65.3	299	1 RL22_DROME	P50887 drosophila
28	32	65.3	124	1 RL22_GADMO	P52865 gadus morhu
29	32	65.3	127	1 RL22_HUMAN	P35268 homo sapien
30	32	65.3	127	1 RL22_MOUSE	P41104 mus musculu
31	32	65.3	127	1 RL22_RAT	P47198 rattus norv
32	32	65.3	130	1 RL22_TRIGR	P13732 tripneustes
33	32	65.3	127	1 RL22_XENLA	P50886 xenopus lae
34	32	65.3	1024	1 RPOB_PLAFA	P21421 plasmodium
35	32	65.3	806	1 SYFB_MYCGE	P24437 mycoplasma
36	32	65.3	484	1 UL32_HSV6U	P52463 herpes simp
37	32	65.3	722	1 VGLH_GPCMV	P87730 guinea pig
38	32	65.3	2376	1 YIM9_YEAST	P40468 saccharomyc
39	31	63.3	214	1 AANT_HDVIT	P29997 hepatitis d
40	31	63.3	205	1 AANT_HDVWO	P29997 hepatitis d
41	31	63.3	502	1 C932_SOYBN	Q42799 glycine max
42	31	63.3	510	1 C933_SOYBN	O81973 glycine max
43	31	63.3	499	1 DLDH_YEAST	P09624 saccharomyc

ALIGNMENTS

RESULT 1	
CC16_YEAST	
ID	CC16_YEAST STANDARD; PRT: 840 AA.
AC	P09798;
DT	01-MAR-1989 (REL. 10, CREATED)
DT	01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE	CELL DIVISION CONTROL PROTEIN 16.
GN	CDC16 OR YKL022C.
OS	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC	EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC	SACCHAROMYCETACEAE; SACCHAROMYCES.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=R.B.WICKNER 1385;
RX	MEDLINE; 88040465.
RA	ICHO T.; WICKNER R.B.;
RT	"Metal-binding, nucleic acid-binding finger sequences in the CDC16
RT	gene of Saccharomyces cerevisiae.";
RL	NUCLEIC ACIDS RES. 15:8439-8450(1987).
RN	[2]
RP	SEQUENCE FROM N.A.
RA	RIEGER M.;
RC	SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN	[3]
RP	TPR REPEATS.
RX	MEDLINE; 90124639.
RA	SIKORSKI R.S., BOGUSKI M.S., GOEBL M., HIETER P.A.;
RT	"A repeating amino acid motif in CDC23 defines a family of proteins
RT	and a new relationship among genes required for mitosis and RNA
RT	synthesis.";
RL	CELL 60:307-317(1990).
RN	[4]
RP	SUBUNITS.
RX	MEDLINE; 95009933.
RA	LAMB J.R., MICHAUD W.A., SIKORSKI R.S., HIETER P.A.;
RT	"Cdc16p, Cdc23p and Cdc27p form a complex essential for mitosis.";
RL	EMBO J. 13:4321-4328(1994).
CC	-1- FUNCTION: EXACT FUNCTION NOT KNOWN. REQUIRED FOR CHROMOSOME
CC	SEGREGATION. MUTATIONS IN CDC16 CAUSE CELLS TO ARREST UNIFORMLY
CC	AT G2/M AFTER DNA REPLICATION BUT PRIOR TO MITOTIC SPINDLE
CC	ELONGATION.
CC	-1- SUBUNIT: CDC16, CDC23 AND CDC27 FORMS A MACROMOLECULAR COMPLEX.
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.
CC	-1- SIMILARITY: CONTAINS 10 TPR DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X06165; G3478; ..
DR	EMBL; Z28022; G486015; ..
DR	PIR; A27832; A27832.
DR	SGD; L0000256; CDC16.
DR	PFAM; PF00515; TPR; 7.
KW	CELL DIVISION; CELL CYCLE; MITOSIS; REPEAT; TPR DOMAIN;
FT	NUCLEAR PROTEIN.
FT	REPEAT 263 295 TPR 1.
FT	REPEAT 296 329 TPR 2.
FT	REPEAT 329 352 TPR 3.
FT	REPEAT 497 530 TPR 4.
FT	REPEAT 531 564 TPR 5.

```

FT REPEAT      565      598      TPR 6.
FT REPEAT      599      632      TPR 7.
FT REPEAT      633      666      TPR 8.
FT REPEAT      674      707      TPR 9.
FT REPEAT      708      741      TPR 10.
SQ SEQUENCE    840 AA; 94991 MW; DFCDE9DA CRC32;

Query Match
Best Local Similarity 71.4%; Score 35; DB 1; Length 840;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKFLKK 9
   :| |||||
Db 650 VKAKYLKK 658

RESULT 2
ID DP3A_BORBU STANDARD; PRT; 1161 AA.
AC OS1526;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA POLYMERASE III, ALPHA CHAIN (EC 2.7.7.7).
GN DNAE OR BB0579.
OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi";
RL NATURE 390:580-586(1997).
CC -!- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -!- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
CC CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE
CC DIMERIZATION TO FORM THE POLIII' COMPLEX. POLIII' ASSOCIATES WITH
CC THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
CC AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE001159; G2688500;
DR TIGR; BB0579;
KW TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION.
SQ SEQUENCE 1161 AA; 132366 MW; 7EBE8AF CRC32;

Query Match
Best Local Similarity 71.4%; Score 35; DB 1; Length 1161;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKFLKK 9
   :| |||||
Db 650 VKAKYLKK 658

Query Match
Best Local Similarity 71.4%; Score 35; DB 1; Length 1161;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 KTKKFLKK 9
   || |||||
Db 89 KTSKFLKK 96

RESULT 3
RL22_SCHPO STANDARD; PRT; 115 AA.
ID AC Q09668; O13694;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE 60S RIBOSOMAL PROTEIN L22.
GN RPL22 OR SPAC11E3.15.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 3-115 FROM N.A.
RA KAWAMUKAI M.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: BELONGS TO THE L22E FAMILY OF RIBOSOMAL PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z98595; E334258;
DR EMBL; D86349; G1408515;
KW RIBOSOMAL PROTEIN.
FT CONFLICT 115
SQ SEQUENCE 115 AA; 13064 MW; 8C8821DF CRC32;

Query Match
Best Local Similarity 71.4%; Score 35; DB 1; Length 115;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TKKFLKK 9
   |||||
Db 76 TKKFLKK 82

RESULT 4
CYSP_PLAFA STANDARD; PRT; 569 AA.
ID AC P25805;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE THROPHOZOITE CYSTEINE PROTEINASE PRECURSOR (EC 3.4.22.-) (TCP).
OS PLASMODIUM FALCIPARUM.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDA; PLASMODIUM.
RN [1]
RP SEQUENCE FROM N.A.
RX ROSENTHAL P.J., NELSON R.G.;
RT "Isolation and characterization of a cysteine proteinase gene of
RT Plasmodium falciparum.";
RL MOL. BIOCHEM. PARASITOL. 51:143-152(1992).
CC -!- FUNCTION: PROBABLY DEGRADATES ERYTHROCYTE HEMOGLOBIN.
CC -!- DEVELOPMENTAL STAGE: THROPHOZOITE.

```



CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE  
 CC PAPAIN FAMILY OF THIOL PROTEASES.  
 CC -!- SIMILARITY: STRONGEST SIMILARITY TO CATHEPSIN L.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M81341; G160248; -  
 CC PIR: A45624; A45624.  
 CC PROSITE: PS00139; THIOL\_PROTEASE\_CYS; 1.  
 CC PROSITE: PS00639; THIOL\_PROTEASE\_HIS; 1.  
 CC PROSITE: PS00640; THIOL\_PROTEASE\_ASN; 1.  
 CC PFAM: PF00112; CYS-protease; 2.  
 CC HSP: P07858; ICSB.  
 CC KW HYDROLASE; THIOL PROTEASE; ZMOGEN; GLYCOPROTEIN; SIGNAL.  
 CC FT SIGNAL 1 ? POTENTIAL.  
 CC FT PROPEP 7 332 ACTIVATION PEPTIDE (POTENTIAL).  
 CC FT CHAIN 333 569 THIOPHOZOITE CYSTEINE PROTEINASE.  
 CC FT DOMAIN 64 70 POLY-ASN.  
 CC FT ACT\_SITE 357 357 BY SIMILARITY.  
 CC FT ACT\_SITE 488 488 BY SIMILARITY.  
 CC FT ACT\_SITE 533 533 BY SIMILARITY.  
 CC FT DISULFID 354 395 BY SIMILARITY.  
 CC FT CARBOHYD 58 58 POTENTIAL.  
 CC FT CARBOHYD 98 98 POTENTIAL.  
 CC FT CARBOHYD 121 121 POTENTIAL.  
 CC FT CARBOHYD 127 127 POTENTIAL.  
 CC FT CARBOHYD 479 479 POTENTIAL.  
 CC FT CARBOHYD 487 487 POTENTIAL.  
 CC SQ SEQUENCE 569 AA; 66880 MW; 9B8F0096 CRC32;

Query Match 69.4%; Score 34; DB 1; Length 569;  
 Best Local Similarity 87.5%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 KTKKFLKK 9  
 Db 23 KKKKFLKK 30  
 RESULT 5  
 ID DPOG\_SCHPO STANDARD; PRT; 1018 AA.  
 AC Q12704;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE DNA POLYMERASE GAMMA (EC 2.7.7.7) (MITOCHONDRIAL DNA POLYMERASE  
 DE CATALYTIC SUBUNIT).  
 DE MIPI.  
 GN SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
 OS EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;  
 OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;  
 CC SCHIZOSACCHAROMYCETES.  
 CC [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-SP808;  
 RX MEDLINE: 96084961.  
 RA ROPP P.A., COPELAND W.C.;  
 RT "Characterization of a new DNA polymerase from Schizosaccharomyces  
 RT pombe: a probable homologue of the Saccharomyces cerevisiae DNA  
 RT polymerase gamma.";  
 RL GENE 165:103-107(1995).  
 CC -!- FUNCTION: INVOLVED IN THE REPLICATION OF MITOCHONDRIAL DNA.  
 CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -  
 CC N PYROPHOSPHATE + DNA(N).  
 CC -!- COFACTOR: MAGNESIUM.

CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
 CC -!- IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES: ALPHA, BETA, GAMMA,  
 CC DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR DIFFERENT REACTIONS  
 CC OF DNA SYNTHESIS.  
 CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z47976; E213759; -  
 CC PROSITE: PS00447; DNA\_POLYMERASE\_A; FALSE\_NEG.  
 CC PFAM: PF00476; DNA\_POL\_A; 1.  
 CC KW TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION;  
 CC DNA-BINDING; MITOCHONDRION; MAGNESIUM.  
 CC SQ SEQUENCE 1018 AA; 116042 MW; 807DD45B CRC32;  
 Query Match 69.4%; Score 34; DB 1; Length 1018;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IKTKKFL 7  
 Db 20 IKTKKFL 26  
 RESULT 6  
 ID Y086\_METJA STANDARD; PRT; 372 AA.  
 AC Q57551;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN MJ0086.  
 GN MJ0086.  
 OS METHANOCOCCLUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCELES; METHANOCOCCEACEAE;  
 CC METHANOCOCCLUS.  
 CC [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE: 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESSE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL SCIENCE 273:1058-1073(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U67466; G1498850; -  
 CC TIGR: MJ0086; -  
 CC KW HYPOTHETICAL PROTEIN.  
 CC SQ SEQUENCE 372 AA; 43285 MW; CE600ED6 CRC32;

Query Match 69.4%; Score 34; DB 1; Length 372;  
Best Local Similarity 75.0%; Pred. No. 23;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
DB 209 ETKKFIKK 216

RESULT 7  
ZFY1\_MOUSE STANDARD; PRT; 782 AA.  
AC P10925;  
DT 01-JUL-1989 (REL. 11, CREATED)  
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE ZINC FINGER Y-CHROMOSOMAL PROTEIN 1.  
GN ZFY1 OR ZFY-1.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-TESTIS.  
RX MEDLINE; 89240049.  
RA ASHWORTH A., SWIFT S., AFFARA N.;  
RT "Sequence of cDNA for murine zfy-1, a candidate for Tdy.";  
RL NUCLEIC ACIDS RES. 17:2864-2864(1989).  
CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL ACTIVATOR.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- SIMILARITY: TO OTHER MEMBERS OF THE ZFY/ZFY FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X14382; G55479; -  
CC PIR; S04047; S04047.  
CC MGD; MGI:99212; ZFY1.  
CC PROSITE; PS00028; ZINC\_FINGER\_C2H2; 7.  
CC PFAM; PF00096; zf-C2H2; 13.  
CC HSP; P08048; 5ZNF.  
CC TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;  
CC DNA-BINDING; REPEAT; NUCLEAR PROTEIN; MULTIGENE FAMILY.  
CC DOMAIN 372 382 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
CC FT DOMAIN 403 777 ZINC-FINGERS.  
CC FT ZN\_FING 403 425 C2H2-TYPE.  
CC FT ZN\_FING 434 456 C2H2-TYPE.  
CC FT ZN\_FING 466 488 C2H2-TYPE.  
CC FT ZN\_FING 497 520 C2H2-TYPE.  
CC FT ZN\_FING 526 548 C2H2-TYPE.  
CC FT ZN\_FING 554 577 C2H2-TYPE.  
CC FT ZN\_FING 583 605 C2H2-TYPE.  
CC FT ZN\_FING 611 634 C2H2-TYPE.  
CC FT ZN\_FING 640 662 C2H2-TYPE.  
CC FT ZN\_FING 668 691 C2H2-TYPE.  
CC FT ZN\_FING 697 719 C2H2-TYPE.  
CC FT ZN\_FING 725 748 C2H2-TYPE.  
CC FT ZN\_FING 754 777 C2H2-TYPE.  
CC SEQUENCE 782 AA; 88265 MW; BDF04024 CRC32;

Query Match 69.4%; Score 34; DB 1; Length 782;  
Best Local Similarity 75.0%; Pred. No. 47;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
DB 413 KTKRFLKR 420

RESULT 8  
ZFY2\_MOUSE STANDARD; PRT; 783 AA.  
AC P20662;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE ZINC FINGER Y-CHROMOSOMAL PROTEIN 2.  
GN ZFY2 OR ZFY-2.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89168416.  
RA MARDON G., PAGE D.C.;  
RT "The sex-determining region of the mouse Y chromosome encodes a  
RT protein with a highly acidic domain and 13 zinc fingers.";  
RL CELL 56:765-770(1989).  
CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL ACTIVATOR.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- SIMILARITY: TO OTHER MEMBERS OF THE ZFY/ZFY FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M24401; G598332; -  
CC PIR; A31491; A31491.  
CC MGD; MGI:99213; ZFY2.  
CC PROSITE; PS00028; ZINC\_FINGER\_C2H2; 7.  
CC PFAM; PF00096; zf-C2H2; 13.  
CC HSP; P08048; 5ZNF.  
CC TRANSCRIPTION REGULATION; ACTIVATOR; ZINC-FINGER; METAL-BINDING;  
CC DNA-BINDING; REPEAT; NUCLEAR PROTEIN; MULTIGENE FAMILY.  
CC DOMAIN 372 382 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
CC FT DOMAIN 403 777 ZINC-FINGERS.  
CC FT ZN\_FING 403 425 C2H2-TYPE.  
CC FT ZN\_FING 434 456 C2H2-TYPE.  
CC FT ZN\_FING 466 488 C2H2-TYPE.  
CC FT ZN\_FING 497 520 C2H2-TYPE.  
CC FT ZN\_FING 526 548 C2H2-TYPE.  
CC FT ZN\_FING 554 577 C2H2-TYPE.  
CC FT ZN\_FING 583 605 C2H2-TYPE.  
CC FT ZN\_FING 611 634 C2H2-TYPE.  
CC FT ZN\_FING 640 662 C2H2-TYPE.  
CC FT ZN\_FING 668 691 C2H2-TYPE.  
CC FT ZN\_FING 697 719 C2H2-TYPE.  
CC FT ZN\_FING 725 748 C2H2-TYPE.  
CC FT ZN\_FING 754 777 C2H2-TYPE.  
CC SEQUENCE 783 AA; 88856 MW; 90E3806A CRC32;

Query Match 69.4%; Score 34; DB 1; Length 783;  
Best Local Similarity 75.0%; Pred. No. 48;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
DB 413 KTKRFLKR 420

RESULT 9  
CYLL\_BOVIN STANDARD; PRT; 667 AA.  
ID CYLL\_BOVIN STANDARD; PRT; 667 AA.  
AC P35662;  
DT 01-JUN-1994 (REL. 29, CREATED)

DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE CYLICIN I (MULTIPLE-BAND POLYPEPTIDE I).  
 GN CYLI OR CYL.  
 OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC ARTODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=TESTIS;  
 RX MEDLINE; 93359502.  
 RA HESS H., HEID H., FRANKE W.W.;  
 RT "Molecular characterization of mammalian cylicin, a basic protein of  
 the sperm head cytoskeleton.";  
 RL J. CELL BIOL. 122:1043-1052(1993).  
 CC -!- FUNCTION: POSSIBLE ARCHITECTURAL ROLE DURING SPERMATOGENESIS. MAY  
 BE INVOLVED IN SPERMATID DIFFERENTIATION.  
 CC -!- SUBCELLULAR LOCATION: CALYX; SPERM HEAD CYTOSKELETAL STRUCTURE.  
 CC -!- TISSUE SPECIFICITY: TESTIS.  
 CC -!- DEVELOPMENTAL STAGE: SPECIFIC TO LATE SPERMATOGENESIS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; 222779; G396151; -;  
 DR PIR; S35913; S35913.  
 DR PIR; A40713; A40713.  
 KW CYTOSKELETON; STRUCTURAL PROTEIN; REPEAT; SPERM; SPERMATOGENESIS.  
 FT DOMAIN 287 569 9 APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 287 305 1.  
 FT REPEAT 306 337 2.  
 FT REPEAT 338 368 3.  
 FT REPEAT 369 405 4.  
 FT REPEAT 406 442 5.  
 FT REPEAT 443 475 6.  
 FT REPEAT 476 516 7.  
 FT REPEAT 517 547 8.  
 FT REPEAT 548 569 9.  
 FT DOMAIN 617 667 PRO-RICH.  
 FT SEQUENCE 667 AA; 74817 MW; 1691124A CRC32;  
 SQ SEQUENCE 667 AA; 74817 MW; 1691124A CRC32;  
 Query Match 67.3%; Score 33; DB 1; Length 667;  
 Best Local Similarity 50.0%; Pred. No. 63;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 IKTKKFLKKT 10  
 : : : : :  
 DB 542 VSSKRYLKKT 551  
 RESULT 10  
 HS30\_ONCTS  
 ID HS30\_ONCTS STANDARD; PRT; 214 AA.  
 AC P42931;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE HEAT SHOCK PROTEIN 30 (HSP 30).  
 GN HSP30.  
 OS ONCORHYNCHUS TSCHAWYTSCHIA (CHINOOK SALMON) (KING SALMON).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;  
 OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES;  
 OC SALMONIDAE; ONCORHYNCHUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HARGIS M., GOFF H., HICKEY E., WEBER L.A.;  
 RL SUBMITTED (JAN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -!- FUNCTION: INVOLVED IN STRESS RESISTANCE AND ACTIN ORGANIZATION  
 CC (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; U19370; G625084; -;  
 DR PROSITE; PS01031; HSP20; 1.  
 DR PFAM; PF00011; HSP20; 1.  
 KW HEAT SHOCK.  
 SQ SEQUENCE 214 AA; 24337 MW; CCA215F6 CRC32;  
 Query Match 67.3%; Score 33; DB 1; Length 214;  
 Best Local Similarity 60.0%; Pred. No. 21;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 IKTKKFLKKT 10  
 : : : : :  
 DB 186 VKTPQFLSKT 195  
 RESULT 11  
 KTR3\_YEAST  
 ID KTR3\_YEAST STANDARD; PRT; 404 AA.  
 AC P38130;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE PROBABLE MANNOSYLTRANSFERASE KTR3 (EC 2.4.1.131).  
 GN KTR3 OR YBR205W OR YBR1445.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 GN SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE; 93377417.  
 RA BUSSEAU F., MALLET L., GAILLON L., JACQUET M.;  
 RT "A 12.8 kb segment, on the right arm of chromosome II from  
 Saccharomyces cerevisiae including part of the DURI2 gene, contains  
 five putative new genes.";  
 RL YEAST 9:797-806(1993).  
 CC -!- FUNCTION: POSSIBLE GLYCOSYLTRANSFERASE.  
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (PROBABLE).  
 CC -!- SIMILARITY: BELONGS TO THE KTR FAMILY OF GLYCOSYLTRANSFERASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; 221487; G311682; -;  
 DR EMBL; 236074; G536583; -;  
 DR PIR; S34031; S34031.  
 DR SGD; L0000926; KTR3.  
 KW TRANSFERASE; GLYCOSYLTRANSFERASE; TRANSMEMBRANE; SIGNAL-ANCHOR.  
 FT DOMAIN 1 27 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 28 44 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
 FT DOMAIN 45 404 LUMENAL (POTENTIAL).  
 FT DOMAIN 45 83 STEM REGION (BY SIMILARITY).  
 FT DOMAIN 84 404 CATALYTIC (BY SIMILARITY).  
 FT ACT\_SITE 295 295 NUCLEOPHILE (POTENTIAL).

SQ SEQUENCE 404 AA; 47482 MW; C99E8E30 CRC32;

Query Match 67.3%; Score 33; DB 1; Length 404;  
Best Local Similarity 85.7%; Pred. No. 39;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TKKFLKK 9  
|||||

DB 257 TKKFIKK 263

RESULT 12

ID KTR7\_YEAST STANDARD; PRT; 517 AA.  
AC P40504;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE PROBABLE MANNOsylTRANSFERASE KTR7 (EC 2.4.1.131).  
GN KTR7 OR YIL085C.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.

RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA BARRELL B.G., BADCOCK K., BANKIER A.T., BOWMAN S., BROWN D.,  
RA CHURCHER C.M., CONNOR R., CORSEY T., DEAR S., DEVLIN K., FRASER A.,  
RA GENTLES S., HAMLYN N., HORSNELL T.S., HUNT S., JAGELS K., JONES M.,  
RA LOUIS E., LYE G., MOULE S., MOULE T., ODELL C., PEARSON D.,  
RA RAJANDREAM M.A., RILES L., ROWLEY N., SKELTON J., SMITH V.,  
RA WALSH S.V., WHITEHEAD S.;  
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [2]

RP CHARACTERIZATION.

RX MEDLINE: 97245297.  
RA LUSSIER M., SDICU A.-M., WINNETT E., VO D.H., SHERAFON J.,  
RA DUESTERHOEFF A., STORMS R.K., BUSSEY H.;

RT "Completion of the Saccharomyces cerevisiae genome sequence allows  
RT identification of KTR5, KTR6 and KTR7 and definition of the nine-  
RT membered KRE2/MNT1 mannosyltransferase gene family in this  
RT organism".  
RL YEAST 13:267-274(1997).

CC -!- FUNCTION: POSSIBLE GLYCOSYLTRANSFERASE.

CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (PROBABLE).

CC -!- SIMILARITY: BELONGS TO THE KTR FAMILY OF GLYCOSYLTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z46728; G57129; -

DR EMBL; Z47047; G763261; -

DR SGD; L0004100; KTR7.

KW TRANSFERASE; GLYCOSYLTRANSFERASE; TRANSMEMBRANE; SIGNAL-ANCHOR;

KW GLYCOPROTEIN.

FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 24 44 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).

FT DOMAIN 45 517 LUMENAL (POTENTIAL).

FT DOMAIN 45 517 STEM REGION (BY SIMILARITY).

FT DOMAIN 86 517 CATALYTIC (BY SIMILARITY).

FT ACT\_SITE 367 367 NUCLEOPHILE (POTENTIAL).

FT CARBOHYD 89 89 POTENTIAL.

FT CARBOHYD 144 144 POTENTIAL.

SQ SEQUENCE 517 AA; 61437 MW; B47FAB82 CRC32;

Query Match

67.3%; Score 33; DB 1; Length 517;

Best Local Similarity 75.0%; Pred. No. 49;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 TKKFLKKT 10  
|||||

DB 260 TKSFIRKT 267

RESULT 13

PTA\_BACSU STANDARD; PRT; 322 AA.  
ID PTA\_BACSU

AC P39646;

DT 01-FEB-1995 (REL. 31, CREATED)

DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)

DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE PROBABLE PHOSPHATE ACETYLTRANSFERASE (EC 2.3.1.8)

DE (PHOSPHOTRANSACETYLASE) (VEGETATIVE PROTEIN 43) (VEG43).

GN PTA OR IPA-88D.

OS BACILLUS SUBTILIS.

OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;

OC BACILLUS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE: 95020537.

RA GLASER P., KUNST F., ARNAUD M., COUDART M.P., GONZALES W.,

RA HULLO M.F., IONESCU M., LUBCHINSKY B., MARCELINO L., MOSZER I.,

RA PRESECAN E., SANTANA M., SCHNEIDER E., SCHWEIZER J., VERTES A.,

RA RAPOPORT G., DANCHIN A.;

RT "Bacillus subtilis genome project: cloning and sequencing of the 97

RT kb region from 325 degrees to 333 degrees.";

RL MOL. MICROBIOL. 10:371-384(1993).

RN [2]

RP SEQUENCE OF 1-10.

RC STRAIN=IS58;

RX MEDLINE: 97443988.

RA ANTELMANN H., BERNHARDT J., SCHMID R., MACH H., VOELKER U.,

RA HECKER M.;

RT "First steps from a two-dimensional protein index towards a response-

RT regulation map for Bacillus subtilis.";

RL ELECTROPHORESIS 18:1451-1463(1997).

CC -!- CATALYTIC ACTIVITY: ACETYL-COA + ORTHOPHOSPHATE = COA +

CC ACETYL-PHOSPHATE.

CC -!- SIMILARITY: BELONGS TO THE PHOSPHATE ACETYLTRANSFERASE AND

CC BUTYRYLTRANSFERASE FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X73124; G580883; -

DR EMBL; Z99123; E1186266; -

DR SUBTILIST; BG10634; PTA

KW TRANSFERASE; ACYLTRANSFERASE.

FT INIT\_MET 0

SQ SEQUENCE 322 AA; 34659 MW; ED661D4C CRC32;

Query Match

67.3%; Score 33; DB 1; Length 322;

Best Local Similarity 70.0%; Pred. No. 31;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKFLKKT 10

|||||

DB 139 IKTKGVAKT 148

RESULT 14

Y011\_MYGE

```

ID Y011_MYCGE STANDARD; PRT; 287 AA.
AC P47257; Q49362;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MG011.
GN MG011.
OS MYCOPLASMA GENITALIUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLIICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE; 96026346.
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,
RA FRITCHMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,
RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL SCIENCE 270:397-403(1995).
RN [2]
RP SEQUENCE OF 158-255 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE; 94075230.
RA PETERSON S.N., HU P.-C., BOTT K.F., HUTCHISON C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing."
RL J. BACTERIOL. 175:7918-7930(1993).
CC -1- SIMILARITY: BELONGS TO THE ATP-GRASP SUPERFAMILY; RIMK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39679; G1045679; -
DR EMBL; U02257; G406920; -
DR TIGR; MG011; -
KW HYPOTHETICAL PROTEIN.
FT CONFLICT 158 158 T -> P (IN REF. 2).
FT CONFLICT 255 255 A -> R (IN REF. 2).
SQ SEQUENCE 287 AA; 33433 MW; 9E2DA74F CRC32;

Query Match 67.3%; Score 33; DB 1; Length 287;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KTKFLKK 9
Db 276 KFKFLKK 283

RESULT 15
Y016_METJA
ID Y016_METJA STANDARD; PRT; 58 AA.
AC Q60322;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL IS-LIKE ELEMENT PROTEIN MJ0016/MJ1621.1.
GN MJ0016 AND MJ1621.1.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;

```

```

RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KIENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT Jannaschii."
RL SCIENCE 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67460; G1498774; -
DR EMBL; U67602; G2826441; -
DR TIGR; MJ0016; -
DR TIGR; MJ1621.1; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 58 AA; 6622 MW; C65DECDB CRC32;

Query Match 67.3%; Score 33; DB 1; Length 58;
Best Local Similarity 60.0%; Pred. No. 6;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 IKTKFLKK 10
Db 9 IKTKFLVK 18

Search completed: September 7, 1999, 23:59:16
Job time: 528 sec

```

**THIS PAGE BLANK (US)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:46 ; Search time 160.2 Seconds  
(without alignments)  
3.842 Million cell updates/sec

Title: US-09-124-280A-36

Perfect score: 49

Sequence: 1 IKTKFLKKT 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	73.5	298	2	O87233 lactococcus
2	36	73.5	239	11	O63545 rattus norv
3	35	71.4	251	1	P94944 methanopyru
4	35	71.4	849	5	Q19070 caenorhabdi
5	35	71.4	974	5	Q23463 caenorhabdi
6	34.5	70.4	379	2	O68797 campylobact
7	34	69.4	1193	2	P71107 clostridium
8	34	69.4	1193	2	O45914 clostridium
9	34	69.4	254	2	O06031 lactococcus
10	34	69.4	2022	4	Q13459 homo sapien
11	34	69.4	1368	5	Q25661 plasmodium
12	34	69.4	403	11	O70533 tokudala os
13	34	69.4	1980	11	O63358 rattus norv
14	34	69.4	207	12	O39628 cydia pomon
15	34	69.4	522	13	Q91GX4
16	33	67.3	219	1	O59032 pyrococcus
17	33	67.3	1104	2	O51934 thermotoga
18	33	67.3	226	2	O67515 aquifex aeo
19	33	67.3	1596	2	O59167 acetobacter
20	33	67.3	202	2	O84499 chlamydia t
21	33	67.3	315	2	O87183 streptococc
22	33	67.3	318	2	O66411 aquifex aeo
23	33	67.3	404	5	O16752 caenorhabdi
24	33	67.3	677	5	Q17973 caenorhabdi
25	33	67.3	134	5	O62226 caenorhabdi
26	33	67.3	650	5	O96580 spissula sol
27	33	67.3	852	10	Q9ZSC8
28	33	67.3	561	11	P70166 mus musculu
29	33	67.3	568	13	Q91572 xenopus lae

#### ALIGNMENTS

```

RESULT 1
O87233          PRELIMINARY;      PRT;   298 AA.
AC O87233;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE POTENTIAL ABC TRANSPORTER.
OS Lactococcus lactis.
OC Plasmid pMRC01.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DPC3147;
RX MEDLINE; 99000510.
RA DOUGHERTY B.A., HILL C., WEIDMAN J.F., RICHARDSON D.R., VENTER J.C.,
RA ROSS R.P.;
RT "Sequence and analysis of the 60 kb conjugative,
RT bacteriocin-producing plasmid pMRC01 from Lactococcus lactis
RT DPC3147."
RL Mol. Microbiol. 29:1029-1038(1998).
DR EMBL; AE001272; AAC56009.1; -.
DR PFAM; PF00005; ABC_tran; 1.
KW Plasmid.
SQ SEQUENCE 298 AA; 33882 MW; 7AE75BEF CRC32;

Query Match          73.5%; Score 36; DB 2; Length 298;
Best Local Similarity 77.8%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKTKFLKK 9
   |||||
Db 276 IKLKKFIKK 284

RESULT 2
O63545          PRELIMINARY;      PRT;   239 AA.
ID O63545;
AC O63545; Q62903;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE DECIDUAL PROLACTIN-RELATED PROTEIN (DPRP).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=MATURE PREGNANT UTERINE DECIDUA;
RX MEDLINE; 93155148.

```

O93386 brachydanio  
 Q9YGX5 brachydanio  
 O28529 archaeglob  
 Q48680 listeria iv  
 O51546 borrelia bu  
 O66612 aquifex aeo  
 O84260 chlamydia t  
 Q9ZDEL rickettsia  
 Q9Z9G2 chlamydia p  
 Q47772 enterococcu  
 Q15029 homo sapien  
 Q23131 caenorhabdi  
 O81617 arabidopsis  
 O08810 mus musculu  
 O11428 avian adeno  
 Q69059 human herpe

RA ROBY K.F., DEB S., GIBORI G., SZPIRER C., LEVAN G., KWOK S.C.,  
 RT SOARES M.J.;  
 RT "Decidual prolactin-related protein. Identification, molecular  
 RT cloning, and characterization."  
 RL J. Biol. Chem. 268:3136-3142(1993).  
 RN [2]  
 RN SEQUENCE OF 1-9 FROM N.A.  
 RC STRAIN-WISTAR; TISSUE=LIVER;  
 RA ORWIG K.O., DAI G., SOARES M.J.;  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; L06441; AAA18219.1; -  
 DR EMBL; U44438; AAB17697.1; -  
 DR PFAM; PF00103; hormone.1.  
 SQ SEQUENCE 239 AA; 27936 MW; 8C552918 CRC32;

Query Match 73.5%; Score 36; DB 11; Length 239;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKFLK 8  
 |||||  
 Db 116 IKTKYLK 123

RESULT 3  
 P94944  
 ID P94944 PRELIMINARY; PRT; 251 AA.  
 AC P94944;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
 DE FWUG, FWUD AND FWUB GENES.  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyrales; Methanopyrus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AV19/DSM 6324;  
 RX MEDLINE; 97231344.  
 RA VORHOUT J.A., VAUPEL M., THAUER R.K.;  
 RT "A selenium-dependent and a selenium-independent formylmethanofuran  
 RT dehydrogenase and their transcriptional regulation in the  
 RT hyperthermophilic Methanopyrus kandleri."  
 RL Mol. Microbiol. 23:1033-1042(1997).  
 DR EMBL; X98917; CAA67413.1; -  
 SQ SEQUENCE 251 AA; 27238 MW; 6CBE84FD CRC32;

Query Match 71.4%; Score 35; DB 1; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLK 8  
 |||||  
 Db 99 KTKFLK 105

RESULT 4  
 Q19070  
 ID Q19070 PRELIMINARY; PRT; 849 AA.  
 AC Q19070;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE ELONGATION FACTOR 2-LIKE (EFT-1).  
 GN EFT-1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 RN Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 92153310.

RA OFULUE E.N., CANDIDO E.P.;  
 RT "Isolation and characterization of eft-1, an elongation factor 2-like  
 RT gene on chromosome III of Caenorhabditis elegans."  
 RL DNA Cell Biol. 11:71-82(1992).  
 DR EMBL; M86958; AAA21824.1; -  
 DR PFAM; PF00679; EFG\_C; 1.  
 DR PFAM; PF00009; GTP\_EFTU; 1.  
 KW Elongation factor.  
 SQ SEQUENCE 849 AA; 96233 MW; 4B2341E5 CRC32;

Query Match 71.4%; Score 35; DB 5; Length 849;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLK 10  
 |||||  
 Db 231 KTRKFKKS 239

RESULT 5  
 Q23463  
 ID Q23463 PRELIMINARY; PRT; 974 AA.  
 AC Q23463;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE THE ABOVE GENBANK ENTRY BEGINS AT AA 120.  
 GN EFT-1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA JONES M., KERSHAW J., HAWKINS J., HILLIER L., JIER M., JOHNSTON L.,  
 RA LIGHNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans."  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA FAVELLO A.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U50193; AAA91248.1; -  
 DR PFAM; PF00679; EFG\_C; 1.  
 DR PFAM; PF00009; GTP\_EFTU; 1.  
 SQ SEQUENCE 974 AA; 110484 MW; ADC84740 CRC32;

Query Match 71.4%; Score 35; DB 5; Length 974;  
 Best Local Similarity 66.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLK 10  
 |||||  
 Db 356 KTRKFKKS 364

RESULT 6  
 O68797  
 ID O68797 PRELIMINARY; PRT; 379 AA.  
 AC O68797;



DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE HEAT SHOCK PROTEIN.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 OC Campylobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA BRAS A.M., WREN B.W., KETLEY J.M.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO OTHER PROKARYOTIC DNAs, AND TO EUKARYOTIC  
 CC DNAs-LIKE PROTEINS  
 CC EMBL: AF053962; AAC08023.1; -;  
 DR PFAM: PF00226; DnaJ; 1.  
 DR PFAM: PF00684; DnaJ\_CXXCXGXG; 1.  
 DR PROSITE: PS00636; DnaJ\_1; 1.  
 DR PROSITE: PS00637; DnaJ\_CXXCXGXG; 1.  
 KW Heat shock; Chaperone; DNA replication.  
 SQ SEQUENCE 379 AA; 42148 MW; EE2BADDI CRC32;

Query Match 70.4%; Score 34.5; DB 2; Length 379;  
 Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
 Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKTKKFLKKT 10  
 ||||| |||||  
 DB 357 IKTKKDFLKKS 367

RESULT 7  
 ID P71107 PRELIMINARY; PRT; 1193 AA.  
 AC P71107;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE A NTNH AND BONT GENES.  
 GN NTNH.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Clostridium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-62A;  
 RX MEDLINE: 97016817.  
 RA EAST A.K., BHANDARI M., STACEY J.M., CAMPBELL K.D., COLLINS M.D.;  
 RT "Organization and phylogenetic interrelationships of genes encoding  
 RT components of the botulinum toxin complex in proteolytic Clostridium  
 RT botulinum types A, B, and F: evidence of chimeric sequences in the  
 RT gene encoding the nontoxic nonhemagglutinin component.";  
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).  
 DR EMBL: X92973; CAA63550.1; -;  
 SQ SEQUENCE 1193 AA; 138019 MW; 21FD4C54 CRC32;

Query Match 69.4%; Score 34; DB 2; Length 1193;  
 Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKK 9  
 ||||| |||||  
 DB 778 IKTKKEFIQK 786

RESULT 8  
 ID Q45914 PRELIMINARY; PRT; 1193 AA.  
 AC Q45914;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE TYPE A PROGENITOR TOXIN NONTXIC-NONHA.  
 GN ANT.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Clostridium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TYPE A NIH;  
 RX MEDLINE: 96096783.  
 RA FUJITA R., FUJINAGA Y., INOUE K., NAKAJIMA H., KUMON H., OGUMA K.;  
 RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin  
 RT components of Clostridium botulinum type A progenitor toxins.";  
 RL FEBS Lett. 376:41-44(1995).  
 RN [2]  
 RP SEQUENCE OF 1-160 FROM N.A.  
 RC STRAIN-TYPE A NIH;  
 RA INOUE K., FUJINAGA Y., WATANABE T., OHYAMA T., TAKESHI K.,  
 RA MORIISHI K., OGUMA K.;  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: D67030; BA11050.1; -;  
 SQ SEQUENCE 1193 AA; 138092 MW; DCBEFFD4 CRC32;

Query Match 69.4%; Score 34; DB 2; Length 1193;  
 Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFLKK 9  
 ||||| |||||  
 DB 778 IKTKKEFIQK 786

RESULT 9  
 ID O06031 PRELIMINARY; PRT; 254 AA.  
 AC O06031;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE EPSC.  
 GN EPSC.  
 OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).  
 OC Plasmid pNZ4000.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Lactococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 97303093.  
 RA VAN KRANENBURG R., MARUGG J.D., VAN SWAM I.I., WILLEM N.J., VOS W.M.;  
 RT "Molecular characterization of the plasmid-encoded eps gene cluster  
 RT essential for exopolysaccharide biosynthesis in Lactococcus lactis.";  
 RL Mol. Microbiol. 24:387-397(1997).  
 DR EMBL: U93384; AAC45230.1; -;  
 KW Plasmid.  
 SQ SEQUENCE 254 AA; 28415 MW; E52B42C2 CRC32;

Query Match 69.4%; Score 34; DB 2; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKTKKFL 7  
 ||||| |||||  
 DB 245 IKTKKFL 251

RESULT 10  
 ID Q13459 PRELIMINARY; PRT; 2022 AA.  
 AC Q13459;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE MYOSIN-IXB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97063843.  
 RA WIRTH J.A., JENSEN K.A., POST P.L., BEMENT W.M., MOOSEKER M.S.;  
 RT "Human myosin-IXb, an unconventional myosin with a chimerin-like  
 rho/rac GTPase-activating protein domain in its tail.";  
 RL J. Cell Sci. 109:653-661(1996).  
 DR EMBL; U42391; AAC50402.1; -.  
 DR PFAM; PF00130; DAG\_PE-bind; 1.  
 DR PFAM; PF00612; IQ; 4.  
 DR PFAM; PF00063; myosin\_head; 2.  
 DR PFAM; PF00788; RA; 1.  
 DR PFAM; PF00620; RhoGAP; 1.  
 SQ SEQUENCE 2022 AA; 228623 MW; E94DD91C CRC32;

Query Match 69.4%; Score 34; DB 4; Length 2022;  
 Best Local Similarity 77.8%; Pred. No. 5.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKKT 10  
 ||| |||:|  
 Db 935 KTKVFLKET 943

RESULT 11  
 Q25661 PRELIMINARY; PRT; 1368 AA.  
 ID Q25661 AC  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
 DE PUTATIVE GLUTAMATE AND ASPARAGINE RICH PROTEIN (FRAGMENT).  
 OS Plasmodium chabaudi.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA WISER M.F., GIRALDO L.E.;  
 RL Submitted (AUG-1995), to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U35149; AAA81320.1; -.  
 FT NON\_TER 1368 1368  
 SQ SEQUENCE 1368 AA; 156529 MW; 39A244DC CRC32;

Query Match 69.4%; Score 34; DB 5; Length 1368;  
 Best Local Similarity 87.5%; Pred. No. 3.9e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TKKFLKKT 10  
 ||| |||:  
 Db 1337 TKKMLKKT 1344

RESULT 12  
 O70533 PRELIMINARY; PRT; 403 AA.  
 ID O70533 AC  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE 2FY (FRAGMENT).  
 OS Tokudaia osimensis.  
 GN Tokudaia osimensis.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Tokudaia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-2N=45;  
 RA SUTOU S.;

RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D83489; BAA48277.1; -.  
 DR PFAM; PF00096; zf-C2H2; 13.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 7.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 403 AA; 47355 MW; DD49750E CRC32;

Query Match 69.4%; Score 34; DB 11; Length 403;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
 ||| |||:  
 Db 33 KTKKFLKK 40

RESULT 13  
 O63358 PRELIMINARY; PRT; 1980 AA.  
 ID Q63358 AC  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE MYOSIN HEAVY CHAIN.  
 GN MYR5.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RX MEDLINE; 95188874.  
 RA REINHARD J., SCHEEL A.A., DIEKMANN D., HALL A., RUPPERT C.,  
 RA BAEHLER M.;  
 RT "A novel type of myosin implicated in signalling by rho family  
 GTPases.";  
 RT GTPases.;  
 RL EMBO J. 14:697-704(1995).  
 DR EMBL; X77609; CAA54700.1; -.  
 DR PFAM; PF00130; DAG\_PE-bind; 1.  
 DR PFAM; PF00612; IQ; 4.  
 DR PFAM; PF00063; myosin\_head; 2.  
 DR PFAM; PF00788; RA; 1.  
 DR PFAM; PF00620; RhoGAP; 1.  
 KW Myosin.  
 SQ SEQUENCE 1980 AA; 225035 MW; 8D89E15C CRC32;

Query Match 69.4%; Score 34; DB 11; Length 1980;  
 Best Local Similarity 77.8%; Pred. No. 5.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKKT 10  
 ||| |||:  
 Db 936 KTKVFLKET 944

RESULT 14  
 O39628 PRELIMINARY; PRT; 207 AA.  
 ID O39628 AC  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last annotation update)  
 DE ORF36L.  
 OS Cydia pomonella granulosis virus (CpGV)  
 OS (Cydia pomonella granulosis virus).  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC JEHLE J.A., VAN DER LINDEN I.F.A., VLAK J.M.;  
 RL Virus Res. 0:0-0(1997).

DR EMBL: AF002732: AAB64424.1; -. D2F0E988 CRC32;  
SQ SEQUENCE 207 AA; 23915 MW; 0; Mismatches 1; Indels 0; Gaps 0;

Query Match 69.4%; Score 34; DB 12; Length 207;  
Best Local Similarity 87.5%; Pred. No. 80;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTKKFLKK 9  
|| |||||  
Db 82 KTHKFLKK 89

## RESULT 15

Q9YGX4 PRELIMINARY; PRT; 522 AA.  
AC Q9YGX4;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE ZOR-2.  
GN ZOR-2.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii;  
OC Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae;  
OC Cyprinidae; Rasbora; Danio.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA SUZUKI H., MAEGAWA S., MURAKAWA M., HOSHIIJIMA K., SHIMURA Y.,  
RA YASUDA K., INOUE K.;  
RT "Identification of zebrafish maternal RNA-binding proteins, ZOR-1 and  
ZOR-2";  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB011681; BAA75638.1; -.  
SQ SEQUENCE 522 AA; 58259 MW; 1CB4E09D CRC32;

Query Match 69.4%; Score 34; DB 13; Length 522;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKTKKFLKK 9  
||| |||||  
Db 451 IKTSKFTKK 459

Search completed: September 7, 1999, 22:47:47  
Job time: 7967 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:07 ; Search time 135.78 Seconds  
(without alignments)  
1.919 Million cell updates/sec

Title: US-09-124-280A-37

Perfect score: 54

Sequence: 1 IKFLKFLKFLK 11

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	11	R33532	Peptide for treati
2	54	100.0	11	R39294	Endotoxin lipid A
3	54	100.0	11	W21625	Antibiotic potenti
4	52	96.3	32	R33514	Generic peptide fo
5	50	92.6	10	R71779	Peptide neutralisi
6	50	92.6	12	R71791	Peptide neutralisi
7	50	92.6	10	W21596	Antibiotic potenti
8	50	92.6	12	W21608	Antibiotic potenti
9	49	90.7	10	W21629	Antibiotic potenti
10	38	70.4	10	R71781	Peptide neutralisi
11	38	70.4	10	W21598	Antibiotic potenti
12	36	66.7	420	W19781	Tyrosyl-tRNA synth
13	35	64.8	7	R33533	Peptide for treati
14	35	64.8	7	R39295	Endotoxin lipid A
15	35	64.8	7	R71778	Peptide neutralisi
16	35	64.8	7	W21626	Antibiotic potenti
17	35	64.8	7	W21595	Antibiotic potenti
18	35	64.8	1279	W98761	H. pylori GHPD 109
19	33	61.1	9	W21628	Antibiotic potenti
20	32	59.3	1077	R91597	C3G protein. Ras p
21	32	59.3	1580	W57412	Homo sapiens sulph
22	31	57.4	2408	R24306	Translation of ORF
23	31	57.4	17	R21386	Sequence of amphip
24	31	57.4	17	R22870	Amphiphilic peptid
25	31	57.4	869	R53732	S. cerevisiae P1cl
26	31	57.4	540	W20605	H. pylori cytoplas
27	31	57.4	566	W20945	H. pylori inner me
28	31	57.4	555	W84154	Human desaturase e
29	31	57.4	608	W84155	Human desaturase e
30	31	57.4	746	W84156	Human desaturase e
31	31	57.4	746	W85135	A desaturase enzym
32	31	57.4	555	W85133	A desaturase enzym
33	31	57.4	608	W85134	A desaturase enzym
34	31	57.4	562	W95512	Amino acid sequenc
35	31	57.4	615	W95513	Amino acid sequenc
36	31	57.4	753	W95514	Amino acid sequenc
37	31	57.4	250	W98373	H. pylori GHPD 110
38	30	55.6	2185	R12141	Enteroviral polype
39	30	55.6	6	R33515	Peptide for treati
40	30	55.6	1029	R38962	GC-A. New purified
41	30	55.6	185	R47115	Toxoplasma GP28.5
42	30	55.6	404	R48674	Putative peach pol
43	30	55.6	287	W06967	Cyclopropane mycol

#### ALIGNMENTS

RESULT 1

ID R33532 standard; peptide; 11 AA.  
AC R33532;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 10; Page 32; 39pp; English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 1; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00038;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11

Db 1 IKFLKFLKFLK 11

RESULT 2

ID R39294 standard; peptide; 11 AA.

AC R39294;

DT 22-DEC-1993 (first entry)

DE Endotoxin lipid A neutralising peptide.

KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;

KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;

KW removal; sera; vaccines; drug solns.; solutions; cytokine release;

KW control; in vivo; in vitro; detoxification; detection;

OS Quantificatio.

PN WO9314115-A.

PD 22-JUL-1993.

PF 14-MAY-1992; E01060.

PR 16-JAN-1992; US-819893.

PA (PORR/) PORRO M.

44 30 55.6 470 1 W12678 DAX-1 protein. New  
45 30 55.6 252 1 W18315 Toxoplasma gondii

PI Porro M;  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 10; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug soles, etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00038;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
 |||||  
 DB 1 IKFLKFLKFLK 11

## RESULT 3

W21625  
 ID W21625 standard; peptide; 11 AA.  
 AC W21625;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #37.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 41; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00038;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
 |||||  
 DB 1 IKFLKFLKFLK 11

## RESULT 4

R33514  
 ID R33514 standard; peptide; 32 AA.  
 AC R33514;

DT 07-JUL-1993 (first entry)  
 DE Generic peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 2; Page 32; 39pp; English.  
 CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X  
 CC where n = 1-10, the peptide is useful for treating or preventing  
 CC septic shock, mixing with polymyxin B to reduce its toxicity;  
 CC removing endotoxins from blood, sera or other fluids (in vivo  
 CC or in vitro); controlling release of cytokines induced by toxins  
 CC endotoxins; as diagnostic reagents to detect and quantify toxins  
 CC in blood or sera; preparing non-toxic antigenic complexes of  
 CC lipid A or lipopolysaccharide (LPS); and for treating pertussis,  
 CC bacterial meningitis and HIV-related infections. The usual dose is  
 CC 10-100 ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 32 AA;

Query Match 96.3%; Score 52; DB 1; Length 32;  
 Best Local Similarity 90.9%; Pred. No. 0.0024;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
 :|||||  
 DB 4 LKFLKFLKFLK 14

## RESULT 5

R71779  
 ID R71779 standard; peptide; 10 AA.  
 AC R71779;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI: 95-075190/10.

PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 10; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new

```

CC peptides.
SQ Sequence 10 AA;

Query Match 92.6%; Score 50; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11
   |||||
Db 1 KFLKFLKFLK 10

RESULT 6
R71791
ID R71791 standard; peptide; 12 AA.
AC R71791;
DE 01-OCT-1995 (first entry)
KW Peptide neutralising toxicity of Lipid A.
OS Endotoxin; Lipid A; septic shock; lipopolysaccharide.
FH Synthetic.
FH Key Location/Qualifiers
FT disulfide_bond 1..12
PN WO9503327-A.
PD 02-FEB-1995.
PF 21-JUL-1994; E02413.
PR 26-JUL-1993; US-097830.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M;
DR WPI; 95-075190/10.
PT New peptide(s) for neutralising LPS endotoxin - comprising
PT repeating units of a basic aminoacid or basic and hydrophobic
PT amino acids
PS Claim 22; Page 21; 26pp; English.
CC New peptides are claimed which are linear or cyclic peptides which
CC include units of formula: (A)n, where A is the cationic amino acid Lys
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or
CC Trp, and p is 2 or greater.
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic
CC activity. Hence they can be used therapeutically to treat septic shock
CC and also in vitro to detoxify vaccines, drug solutions, injectable
CC nutrient solutions, etc.
CC The present sequence is a specifically claimed example of the new
CC peptides.
SQ Sequence 12 AA;

Query Match 92.6%; Score 50; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11
   |||||
Db 2 KFLKFLKFLK 11

RESULT 7
W21596
ID W21596 standard; peptide; 10 AA.
AC W21596;
DE 26-AUG-1997 (first entry)
KW Antibiotic potentiating peptide #8.
DE Potentiate; antibiotic; microbial infection; lipopolysaccharide;
KW permeability; outer bacterial membrane.
OS Synthetic.
PN WO9638163-A1.
PD 05-DEC-1996.
PF 29-MAY-1996; E02313.
PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M;

```

```

DR WPI; 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 12; Page 24; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to
CC potentiate the activity of an antibiotic when they are co-administered
CC with the antibiotic. Compositions containing these peptides are used
CC to treat or prevent microbial infections. These peptides bind to
CC lipopolysaccharide on the bacteria so may increase permeability of
CC the outer bacterial membrane to the antibiotic, allowing a reduction
CC in the dose of antibiotic required by 10-90% of the normal dose for
CC in vivo or in vitro application. Any toxic side effects are
CC correspondingly reduced.
SQ Sequence 10 AA;

Query Match 92.6%; Score 50; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11
   |||||
Db 1 KFLKFLKFLK 10

RESULT 8
W21608
ID W21608 standard; peptide; 12 AA.
AC W21608;
DE 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #20.
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;
KW permeability; outer bacterial membrane.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide_bond 1..12
PN WO9638163-A1.
PD 05-DEC-1996.
PF 29-MAY-1996; E02313.
PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M;
DR WPI; 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 24; Page 26; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to
CC potentiate the activity of an antibiotic when they are co-administered
CC with the antibiotic. Compositions containing these peptides are used
CC to treat or prevent microbial infections. These peptides bind to
CC lipopolysaccharide on the bacteria so may increase permeability of
CC the outer bacterial membrane to the antibiotic, allowing a reduction
CC in the dose of antibiotic required by 10-90% of the normal dose for
CC in vivo or in vitro application. Any toxic side effects are
CC correspondingly reduced.
SQ Sequence 12 AA;

Query Match 92.6%; Score 50; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11
   |||||
Db 2 KFLKFLKFLK 11

RESULT 9
W21629
ID W21629 standard; peptide; 10 AA.
AC W21629;
DE 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #41.

```

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 45; Page 29; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 90.7%; Score 49; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0026;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
 DB 1 IKFLKFLKFL 10

## RESULT 10

ID R71781 standard; peptide; 10 AA.  
 AC R71781; 70.4%; Score 38; DB 1; Length 10;  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 12; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 70.4%; Score 38; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 0.22;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11

DB 1 KFFKFFKFFK 10

## RESULT 11

ID W21598 standard; peptide; 10 AA.  
 AC W21598;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #10.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 14; Page 25; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 70.4%; Score 38; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 0.22;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 DB 1 KFFKFFKFFK 10

## RESULT 12

ID W19781 standard; Protein; 420 AA.  
 AC W19781;  
 DE Tyrosyl-tRNA synthetase from Staph. aureus.  
 KW Tyrosyl-tRNA synthetase; Bacillus subtilis; tyrs; immunological response;  
 KW antibody; bacterial infection; adherence; damaged tissue; wound healing;  
 KW vaccine; skin; protection.  
 OS Staphylococcus aureus.  
 PN EP-785258-A1.  
 PD 23-JUL-1997.  
 PF 17-JAN-1997; 300307.  
 PR 19-JAN-1996; GB-001067.  
 PA (SMIK) SMITHKLINE BEECHAM PLC.  
 PI Hodgson JE, Lawlor EJ;  
 DR WPI; 97-365933/34.  
 DR N-PSDB; T72807.  
 PT DNA encoding tyrosyl-tRNA synthetase from Staphylococcus aureus WCUH  
 PT 29 - useful for protection against bacterial infections  
 PS Claim 14; Page 19-21; 31pp; English.  
 CC The present sequence is a novel tyrosyl tRNA synthetase protein from  
 CC Staphylococcus aureus strain WCUH29 (NCIMB 40771), which is related by  
 CC amino acid sequence homology to Bacillus subtilis tRNA synthetase encoded  
 CC by the tyrs gene. The enzyme catalyses the aminoacylation of tRNA-Tyr, by  
 CC a two step mechanism. The first step involves formation of a stable  
 CC enzyme-tyrosinyl adenylate complex resulting from the specific binding  
 CC and reaction of ATP and L-tyrosine. Subsequently, the 3' terminal  
 CC adenosine of enzyme-bound tRNA-Tyr reacts with the aminoacyl adenylate,



CC leading to the esterification of the tRNA and release of AMP. Vectors  
 CC comprising the DNA (or polynucleotides having at least 70 % identity to  
 CC it) can be used for the recombinant production of the enzyme. The enzyme  
 CC or its related DNA (through gene therapy) is used to induce an  
 CC immunological response in a mammal to generate antibodies to protect  
 CC against disease. The antibodies protect against invasion of bacteria,  
 CC e.g. by blocking adherence of bacteria to damaged tissue, including  
 CC wounds in skin or connective tissue caused by mechanical, chemical or  
 CC thermal damage or by implantation of in-dwelling devices, or wounds in  
 CC the mucous membranes.  
 SQ Sequence 420 AA;

Query Match 66.7%; Score 36; DB 1; Length 420;  
 Best Local Similarity 70.0%; Pred. No. 18;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKELKFL 10  
 |||||: ||  
 Db 265 IKFLKYFTFL 274

## RESULT 13

R33533  
 ID R33533 standard; peptide; 7 AA.

AC R33533;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (FORR/) PORRO M.  
 PI Porro M;

DR WPI; 93-094304/11.  
 DT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 11: Page 32; 39pp; English.

CC This peptide is a specific example of a generic peptide of  
 CC formula R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 7 AA;

Query Match 64.8%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLK 8  
 |||||  
 Db 1 KFLKFLK 7

## RESULT 14

R39295

ID R39295 standard; peptide; 7 AA.

AC R39295;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (FORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 DT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 11: Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 7 AA;

Query Match 64.8%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLK 8  
 |||||  
 Db 1 KFLKFLK 7

## RESULT 15

R71778  
 ID R71778 standard; peptide; 7 AA.

AC R71778;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.

DT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids

PS Claim 9: Page 20; 26pp; English.

CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, infectable  
 CC nutrient solutions, etc.

CC The present sequence is a specifically claimed example of the new  
CC peptides.

SO Sequence 7 AA;

Query Match 64.88; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLK 8

Db 1 KFLKFLK 7

Search completed: September 7, 1999, 22:50:07  
Job time: 7888 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:50 ; Search time 82.37 Seconds  
(without alignments)  
1.318 Million cell updates/sec

Title: US-09-124-280A-37  
Perfect score: 54  
Sequence: 1 IKFLKFLKFLK 11

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/PCTUS9\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	11	1	US-08-049-871-8
2	54	100.0	11	1	US-07-819-893-8
3	54	100.0	11	1	US-08-280-397-8
4	54	100.0	11	2	US-08-456-112B-37
5	50	92.6	10	1	US-08-097-830E-8
6	50	92.6	12	1	US-08-097-830E-20
7	50	92.6	10	2	US-08-456-112B-8
8	50	92.6	12	2	US-08-456-112B-20
9	49	90.7	10	2	US-08-456-112B-41
10	38	70.4	10	1	US-08-097-830E-10
11	38	70.4	10	2	US-08-456-112B-10
12	35	64.8	7	1	US-07-819-893-9
13	35	64.8	7	1	US-08-280-397-9
14	35	64.8	7	1	US-08-097-830E-7
15	35	64.8	7	2	US-08-456-112B-7
16	35	64.8	7	2	US-08-456-112B-38
17	33	61.1	9	2	US-08-456-112B-40
18	31	57.4	17	1	US-07-725-331-49
19	31	57.4	2254	2	US-08-286-819A-28
20	31	57.4	17	3	PCT-US91-05047-49
21	30	55.6	35	1	US-07-938-782A-10
22	30	55.6	287	1	US-08-457-245-3
23	30	55.6	35	1	US-08-630-524-10
24	30	55.6	185	1	US-08-338-543-2
25	30	55.6	576	2	US-08-576-279-58
26	29	53.7	17	1	US-07-725-331-53
27	29	53.7	802	1	US-08-015-985-1
28	29	53.7	793	1	US-08-015-985-3
29	29	53.7	236	1	US-08-015-985-6
30	29	53.7	699	1	US-08-348-008B-7
31	29	53.7	135	1	US-08-468-853-4
32	29	53.7	1141	1	US-08-363-300-2
33	29	53.7	135	2	US-08-468-855-4
34	29	53.7	627	2	US-08-703-947-2
35	29	53.7	135	2	US-08-310-357-4
36	29	53.7	135	2	US-08-468-852-4
37	29	53.7	238	2	US-08-380-403A-7
38	29	53.7	699	2	US-08-800-825A-7
39	29	53.7	17	3	PCT-US91-05047-53

40 28 51.9 17 1 US-07-725-331-50 Sequence 50, Appl  
41 28 51.9 22 1 US-07-965-663A-19 Sequence 19, Appl  
42 28 51.9 887 1 US-08-215-709-1 Sequence 1, Appl  
43 28 51.9 32 1 US-08-118-270-188 Sequence 188, App  
44 28 51.9 20 1 US-08-311-611A-101 Sequence 101, App  
45 28 51.9 20 3 PCT-US95-00636-101 Sequence 101, App

ALIGNMENTS

RESULT 1  
US-08-049-871-8  
; Sequence 8, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-8

Query Match 100.0%; Score 54; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
|||||  
DB 1 IKFLKFLKFLK 11

RESULT 2  
US-07-819-893-8  
; Sequence 8, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: circular  
US-07-819-893-8

Query Match 100.0%; Score 54; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFLK 11  
| | | | | | | | | |  
Db 1 IKFLKFLKFLK 11

RESULT 3  
US-08-280-397-8  
;; Sequence 8, Application US/08280397  
;; Patent No. 5589459  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
US-08-280-397-8

Query Match 100.0%; Score 54; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFLK 11  
| | | | | | | | | |  
Db 1 IKFLKFLKFLK 11

RESULT 4  
US-08-456-112B-37  
;; Sequence 37, Application US/08456112B  
;; Patent No. 5834430  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 37:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
US-08-456-112B-37

Query Match 100.0%; Score 54; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
| | | | | | | | | |  
Db 1 IKFLKFLKFLK 11

RESULT 5  
US-08-097-830E-8  
; Sequence 8, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-8

Query Match 92.6%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
| | | | | | | | | |  
Db 1 KFLKFLKFLK 10

RESULT 6  
US-08-097-830E-20  
; Sequence 20, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-20

Query Match 92.6%; Score 50; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0049;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
| | | | | | | | | |  
Db 2 KFLKFLKFLK 11

RESULT 7  
US-08-456-112B-8  
; Sequence 8, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-8

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-8

Query Match 92.6%; Score 50; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11  
| | | | |  
Db 1 KFLKFLKFLK 10

## RESULT 8

US-08-456-112B-20

; Sequence 20, Application US/08456112B

; Patent No. 5834430

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson &amp; Costigan

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: LEADING EDGE 486

; OPERATING SYSTEM: DOS

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/456,112B

; FILING DATE: May 31, 1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Costigan, James V.

; REGISTRATION NUMBER: 25,669

; REFERENCE/DOCKET NUMBER: 576-004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 302-8998

; TELEFAX: (212) 302-8998

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12 amino acids

; TYPE: amino acid

; TOPOLOGY: circular

US-08-456-112B-20

Query Match 92.6%; Score 50; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0049;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFLK 11  
| | | | |  
Db 2 KFLKFLKFLK 11

## RESULT 9

US-08-456-112B-41

; Sequence 41, Application US/08456112B

; Patent No. 5834430

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson & Costigan

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: LEADING EDGE 486

; OPERATING SYSTEM: DOS

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/456,112B

; FILING DATE: May 31, 1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Costigan, James V.

; REGISTRATION NUMBER: 25,669

; REFERENCE/DOCKET NUMBER: 576-004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 302-8989

; TELEFAX: (212) 302-8998

; INFORMATION FOR SEQ ID NO: 41:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10 amino acids

; TYPE: amino acid

; TOPOLOGY: circular

US-08-456-112B-41

Query Match 90.7%; Score 49; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0061;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFL 10  
| | | | |  
Db 1 IKFLKFLKFL 10

## RESULT 10

US-08-097-830B-10

; Sequence 10, Application US/08097830E

; Patent No. 5652211

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: Peptides For Neutralizing The

; TITLE OF INVENTION: Toxicity of Lipid A

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson &amp; Costigan, P.C.

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: DOS

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/097,830E

; FILING DATE:

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-10

Query Match 70.4%; Score 38; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 0.37;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFLKFLK 11  
|| || || ||  
Db 1 KFFKFFK 10

RESULT 11  
US-08-456-112B-10  
Sequence 10, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-10

Query Match 70.4%; Score 38; DB 2; Length 10;  
Best Local Similarity 70.0%; Pred. No. 0.37;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KFLKFLK 11  
|| || || ||  
Db 1 KFFKFFK 10

RESULT 12  
US-07-819-893-9  
Sequence 9, Application US/07819893  
Patent No. 5371186  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/819,893  
FILING DATE: 19920115  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
US-07-819-893-9

Query Match 64.8%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLK 8  
|||||  
Db 1 KFLKFLK 7

RESULT 13  
US-08-280-397-9  
Sequence 9, Application US/08280397  
Patent No. 5589459  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,397  
FILING DATE: 07/26/94  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-280-397-9

Query Match 64.8%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLK 8  
Db 1 KFLKFLK 7

RESULT 14  
US-08-097-830E-7  
Sequence 7, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TOXICITY OF LIPID A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-7

Query Match 64.8%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLK 8  
Db 1 KFLKFLK 7

RESULT 15  
US-08-456-112B-7  
Sequence 7, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-7

Query Match 64.8%; Score 35; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLK 8  
Db 1 KFLKFLK 7

Search completed: September 7, 1999, 23:07:50  
Job time: 1748 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:52 ; Search time 105.14 Seconds  
(without alignments)  
4.192 Million cell updates/sec

Title: US-09-124-280A-37  
Perfect score: 54  
Sequence: 1 IKFLKFLKFLK 11

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	70.4	394	2 F69888	alanine racemase h
2	37	68.5	201	2 G70130	conserved hypothet
3	36	66.7	956	2 A45506	H+-transporting AT
4	36	66.7	956	2 A43637	H+-transporting AT
5	36	66.7	956	2 S50751	H+-transporting AT
6	36	66.7	951	2 S52728	H+-transporting AT
7	36	66.7	949	2 T02083	H+-transporting AT
8	35	64.8	979	2 J00894	p115 protein - Myc
9	35	64.8	1279	2 E64709	type IIS restricti
10	35	64.8	1278	2 A71609	probable secreted
11	34.5	63.9	238	1 E64304	probable phosphoe
12	34	63.0	153	2 C64373	hypothetical prote
13	34	63.0	260	2 A69068	2-hydroxyhepta-2,4
14	34	63.0	583	2 S43139	phytoene dehydrog
15	34	63.0	578	2 S55102	PSP2 protein - yea
16	34	63.0	337	2 S53846	ribosomal protein
17	34	63.0	564	2 E70394	ABC transporter (h
18	34	63.0	132	2 C71095	hypothetical prote
19	33	61.1	1858	1 A42114	genome polyprotein
20	33	61.1	138	2 S66004	transcription regu
21	33	61.1	158	2 S61604	probable membrane
22	33	61.1	192	2 H70138	cell division prot
23	33	61.1	2291	2 S11238	polymerase - Berne
24	33	61.1	535	2 A64597	conserved hypothet
25	33	61.1	286	2 B70614	probable mnaal pro
26	33	61.1	140	2 H64411	hypothetical prote
27	33	61.1	4572	2 S57908	hypothetical poly
28	33	61.1	110	2 S55208	hypothetical prote
29	33	61.1	1122	2 S64443	probable membrane
30	33	61.1	1659	2 JC4956	vitellogenin precu
31	33	61.1	141	2 B70485	hypothetical prote
32	33	61.1	644	2 B70420	NADH dehydrogenase
33	32.5	60.2	585	2 S11783	bla regulator prot
34	32.5	60.2	771	2 S51421	hypothetical prote
35	32	59.3	422	1 A42648	tyrosine-tRNA lig
36	32	59.3	1866	1 GNWE2C	genome polyprotein
37	32	59.3	190	2 C39746	NADH dehydrogenase
38	32	59.3	201	2 S62757	NADH dehydrogenase
39	32	59.3	236	2 A56509	peroxisomal membra

40	32	59.3	936	2 B64567	cytochrome c bioge
41	32	59.3	936	2 H71862	probable cytochrom
42	32	59.3	119	2 S15218	vsdf protein - Sal
43	32	59.3	121	2 C71696	hypothetical prote
44	32	59.3	475	2 H71660	nitrogen assimilat
45	32	59.3	1113	3 JE0315	low-density lipopr

ALIGNMENTS

RESULT 1

F69888

alanine racemase homolog yncD - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998

C:Accession: F69888

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabet, C.; Ferrari, Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino A:Authors: Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete Y, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl Rieger, M.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili

A:Reference number: A69580; MUID:98044033

A:Accession: F69888

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-394 <KUN>

A:Cross-references: GB:299113; GB:AL009126; NID:g2634090; PID:e1183423; PID:g2634148

A:Experimental source: strain 168

C:Genetics:

A:Gene: yncD

C:Superfamily: alanine racemase

Query Match 70.4%; Score 38; DB 2; Length 394;

Best Local Similarity 70.0%; Pred. No. 12;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
||: || |||

Db 187 KFIISFLSFLK 196

RESULT 2

G70130

conserved hypothetical protein BB0247 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Dec-1998

C:Accession: G70130

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943

A:Accession: G70130

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-201 <KLE>

A:Cross-references: GB:AE001135; GB:AE000783; NID:g2688144; PID:g2688151; TIGR:BB0247

A:Experimental source: strain B31

C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0226

Query Match 68.5%; Score 37; DB 2; Length 201;  
 Best Local Similarity 80.0%; Pred. No. 9;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
 I I I I I I I  
 Db 185 IAFKFLKFL 194

RESULT 3  
 A45506  
 H+-transporting ATPase (EC 3.6.1.35) LHAI - tomato  
 C:Species: Lycopersicon esculentum (tomato)  
 C:Date: 14-May-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Mar-1998  
 C:Accession: A45506  
 R:Ewing, N.N.; Wimmers, L.E.; Meyer, D.J.; Chetelat, R.T.; Bennett, A.B.  
 Plant Physiol. 94, 1874-1881, 1990  
 A:Title: Molecular cloning of tomato plasma membrane H+-ATPase.  
 A:Reference number: A45506  
 A:Accession: A45506  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-956 <EWI>  
 A:Cross-references: GB:M60166; NID:gl70463; PID:gl70464  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; hydrolase; phosphoprotein; transmembrane protein  
 F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:427/Binding site: ATP (Lys) #status predicted

Query Match 66.7%; Score 36; DB 2; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 I I I I I I I  
 Db 61 KFLKFLGFM 69

RESULT 4  
 A43637  
 H+-transporting ATPase (EC 3.6.1.35) - curled-leaved tobacco  
 C:Species: Nicotiana glauca (curled-leaved tobacco)  
 C:Date: 11-Feb-1993 #sequence\_revision 12-Mar-1993 #text\_change 17-Mar-1999  
 C:Accession: A43637  
 R:Boutry, M.; Michelet, B.; Goffeau, A.  
 Biochem. Biophys. Res. Commun. 162, 567-574, 1989  
 A:Title: Molecular cloning of a family of plant genes encoding a protein homologous to P  
 A:Reference number: A43637; MUID:99334860  
 A:Accession: A43637  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-956 <BOU>  
 A:Cross-references: GB:M27888; NID:gl70205; PID:gl70206  
 A:Note: The authors translated the codon CGG for residue 131 as Trp  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; hydrolase; phosphoprotein  
 F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:427/Binding site: ATP (Lys) #status predicted

Query Match 66.7%; Score 36; DB 2; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 I I I I I I I  
 Db 61 KFLKFLGFM 69

RESULT 5  
 S50751  
 H+-transporting ATPase (EC 3.6.1.35) (clone PHA1) - potato  
 C:Species: Solanum tuberosum (potato)  
 C:Date: 14-Jul-1995 #sequence\_revision 08-Sep-1995 #text\_change 24-Apr-1998  
 C:Accession: S50751; S39313  
 R:Harms, K.; Woehner, R.V.; Schulz, B.; Frommer, W.B.  
 Plant Mol. Biol. 26, 979-988, 1994  
 A:Title: Isolation and characterization of P-type H(+)-ATPase genes from potato.  
 A:Reference number: S50751  
 A:Accession: S50751  
 A:Molecule type: mRNA  
 A:Residues: 1-956 <HAR>  
 A:Cross-references: EMBL:X76536; NID:g435002; PID:g435003  
 A:Note: the nucleotide is not given  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; hydrolase; phosphoprotein  
 F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:427/Binding site: ATP (Lys) #status predicted

Query Match 66.7%; Score 36; DB 2; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 I I I I I I I  
 Db 61 KFLKFLGFM 69

RESULT 6  
 S52728  
 H+-transporting ATPase (EC 3.6.1.35) - kidney bean  
 C:Species: Phaseolus vulgaris (kidney bean)  
 C:Date: 19-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 24-Apr-1998  
 C:Accession: S52728  
 R:Campos, F.; Perez-Castineira, J.R.; Villalba, J.M.; Serrano, R.  
 submitted to the EMBL Data Library, March 1995  
 A:Reference number: S52728  
 A:Accession: S52728  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-951 <CAM>  
 A:Cross-references: EMBL:X85804; NID:g758249; PID:g758250  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
 C:Keywords: ATP; hydrolase; phosphoprotein  
 F:483-655/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:331/Active site: Asp (aspartylphosphate intermediate) #status predicted  
 F:425/Binding site: ATP (Lys) #status predicted

Query Match 66.7%; Score 36; DB 2; Length 951;  
 Best Local Similarity 77.8%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 I I I I I I I  
 Db 59 KFLKFLGFM 67

RESULT 7  
 T02083  
 H+-transporting ATPase (EC 3.6.1.35) Mhal - maize  
 C:Species: Zea mays (maize)  
 C:Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 26-Feb-1999  
 C:Accession: T02083  
 R:Jin, Y.; Bennetzen, J.L.  
 Plant Cell 6, 1177-1186, 1994  
 A:Title: Integration and nonrandom mutation of a plasma membrane proton ATPase gene f  
 A:Reference number: Z14555  
 A:Accession: T02083

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-949 <IIN>

A:Cross-references: EMBL:U09989; NID:G507770; PID:G533775

A:Experimental source: strain D3L

C:Genetics:

A:Gene: Mhal

A:Map position: 2L

A:Introns: 21/3; 61/3; 94/3; 140/3; 185/3; 245/3; 314/3; 354/3; 394/3; 435/3; 470/3; 519

C:Function:

A:Description: transports protons across the plasma membrane to regulate intracellular pH

C:Keywords: ATP; hydrogen ion transport; hydrolase; phosphoprotein

#### Query Match

66.7%; Score 36; DB 2; Length 949;

Best Local Similarity 77.8%; Pred. No. 60;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10

|||||

Db 64 KFLKFLGFM 72

#### RESULT 8

JO0894

Pl15 protein - Mycoplasma hyorhinis (SGC3)

C:Species: Mycoplasma hyorhinis

C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 04-Sep-1998

C:Accession: JO0894

R:Notarnicola, S.M.; McIntosh, M.A.; Wise, K.S.

Gene 97, 77-85, 1991

A:Title: A Mycoplasma hyorhinis protein with sequence similarities to nucleotide-binding

A:Reference number: JO0894; MUID:91138990

A:Molecule type: DNA

A:Residues: 1-979 <NOT>

A:Cross-references: GB:M34956

A:Note: the authors translated the codon AAA for residue 956 as Leu

C:Comment: This protein is located in the cytoplasm.

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: conserved hypothetical Pl15 protein

C:Keywords: P-loop

F:32-39/Region: nucleotide-binding motif A (P-loop)

#### Query Match

64.8%; Score 35; DB 2; Length 979;

Best Local Similarity 54.5%; Pred. No. 91;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11

|||||

Db 918 IRYEFLKLLK 928

#### RESULT 9

E64709

Type IIS restriction enzyme R and M protein - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 10-Oct-1997

C:Accession: E64709

R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: E64709

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1279 <TOM>

A:Cross-references: GB:AE000649; GB:AE000511; NID:G2314687; PID:G2314695; TIGR:HP1517

Query Match 64.8%; Score 35; DB 2; Length 1279;

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11

|||||

Db 738 KELKFKFLK 747

#### RESULT 10

A71609

Probable secreted protein PFB0675w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998

C:Accession: A71609

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.

; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.

Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600

A:Accession: A71609

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1278 <GAR>

A:Cross-references: GB:AE001410; GB:AE001362; NID:G3845245; PID:G3845247; TIGR:PFB067

A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0675w

#### Query Match

64.8%; Score 35; DB 2; Length 1278;

Best Local Similarity 70.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10

|||||

Db 687 IKFKFLFLFI 696

#### RESULT 11

E64304

Probable phosphoesterase (EC 3.1.-.-) MJ0037 - Methanococcus jannaschii

N:Alternate names: hypothetical protein MJ0037

C:Species: Methanococcus jannaschii

C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 24-Jul-1998

C:Accession: E64304

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasch

A:Reference number: A64300; MUID:96337999

A:Accession: E64304

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-238 <BUL>

A:Cross-references: GB:U67462; GB:L77117; NID:G1590835; PID:G1498797; TIGR:MJ0037

C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.

C:Genetics:

A:Map position: FOR38073-38789

C:Superfamily: probable phosphoesterase MJ0037; phosphoesterase core homology

C:Keywords: hydrolase

F:22-109/Domain: phosphoesterase core homology <PEC>

#### Query Match

63.9%; Score 34.5; DB 1; Length 238;

Best Local Similarity 66.7%; Pred. No. 28;

Matches 8; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKFLK-FLKFLK 11

Db 85 IKFLKFLK 96  
||||| :|||:

## RESULT 12

C64373  
hypothetical protein MJ0587 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1998  
C:Accession: C64373  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999  
A:Accession: C64373  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-153 <BUL>  
A:Cross-references: GB:U67507; GB:L77117; NID:G1591288; PID:G1591295; TIGR:MJ0587; PID:G  
C:Genetics:  
A:Map position: REV519996-519535  
C:Superfamily: conserved hypothetical protein MJ0129

Query Match 63.0%; Score 34; DB 2; Length 153;  
Best Local Similarity 63.6%; Pred. No. 22;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLK 11  
|||:| :|||

Db 138 IKYLIIRNLK 148

## RESULT 13

A69068  
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase - Methanobacterium thermoautotrophicum (str  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 13-Sep-1998  
C:Accession: A69068  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514  
A:Accession: A69068  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-260 <WTH>  
A:Cross-references: GB:AF000911; GB:AF000666; NID:G2622623; PID:G2622625  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1507  
A:Start codon: GTG  
C:Superfamily: 2-Hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 63.0%; Score 34; DB 2; Length 260;  
Best Local Similarity 75.0%; Pred. No. 37;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 LKFLKFLK 11  
||||| :|||:

Db 4 LKFLKFLK 11

## RESULT 14

S43139  
phytoene dehydrogenase (EC 1.3.-.-) - Phycomyces blakesleeanus  
N:Alternate names: phytoene desaturase

C:Species: Phycomyces blakesleeanus  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Sep-1997  
C:Accession: S43139  
R:Ruiz-Hidalgo, M.J.  
submitted to the EMBL Data Library, March 1994  
A:Reference number: S43139  
A:Accession: S43139  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-583 <RUI>  
A:Cross-references: EMBL:X78434; NID:G468556; PID:G468557  
C:Genetics:  
A:Introns: 197/2  
C:Keywords: oxidoreductase

Query Match 63.0%; Score 34; DB 2; Length 583;  
Best Local Similarity 66.7%; Pred. No. 82;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 FLKFLKFLK 11  
|||:| :|||

Db 122 FLRLDFMK 130

## RESULT 15

S55102  
PSP2 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YN9571.01; protein YML017w  
C:Species: Saccharomyces cerevisiae  
C>Date: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 06-Feb-1998  
C:Accession: S55102; S59358  
R:Gentles, S.; Bowman, S.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: S55102  
A:Accession: S55102  
A:Molecule type: DNA  
A:Residues: 1-578 <GEN>  
A:Cross-references: EMBL:Z49810; NID:G854472; PID:G854473; MIPS:YML017w  
A:Experimental source: strain AB972  
R:Formosa, T.D.; Nittis, I.  
submitted to the EMBL Data Library, August 1995  
A:Description: High copy suppressors of the temperature sensitivity of DNA polymerase  
A:Reference number: S59358  
A:Accession: S59358  
A:Molecule type: DNA  
A:Residues: 1-578 <FOR>  
A:Cross-references: EMBL:U33116; NID:G995752; PID:G995753  
C:Genetics:  
A:Gene: SGD:PSP2  
A:Cross-references: SGD:S0004479; MIPS:YML017w  
A:Map position: 13L

Query Match 63.0%; Score 34; DB 2; Length 578;  
Best Local Similarity 75.0%; Pred. No. 81;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLK 9  
|||:| :|||

Db 117 KFTKFKF 124

Search completed: September 7, 1999, 23:22:55  
Job time: 962 sec